



**DIVERSIDADE GENÉTICA E RESISTÊNCIA NATURAL AO
MARAVIROC EM ESTIRPES DO VÍRUS DA
IMUNODEFICIÊNCIA HUMANA TIPO 1 (HIV-1) EM
CIRCULAÇÃO EM UTILIZADORES DE DROGAS POR VIA
ENDOVENOSA NA GRANDE LISBOA**

CARINA ISABEL CORREIA DE SOUSA

**DISSERTAÇÃO PARA A OBTENÇÃO DO GRAU DE
MESTRE EM MICROBIOLOGIA MÉDICA**

SETEMBRO 2012



**Diversidade genética e resistência natural ao maraviroc em estirpes do
vírus da imunodeficiência humana tipo 1 (HIV-1) em circulação em
utilizadores de drogas por via endovenosa na Grande Lisboa**

Carina Isabel Correia de Sousa

Dissertação para obtenção do grau de Mestre em Microbiologia Médica

Orientador: Professor Doutor João Piedade

Unidade de Microbiologia Médica - Grupo de Virologia,
Instituto de Higiene e Medicina Tropical,
Universidade Nova de Lisboa

Setembro 2012

Elementos Bibliográficos

Carina Sousa, Sandra Videira e Castro, Teresa Venenno, Elizabeth Pádua, Aida Esteves, Ricardo Parreira, João Piedade. Genetic diversity and natural resistance to Maraviroc in HIV-1 strains circulating among intravenous drug users in the Greater Lisbon. Microbiotec'11, Braga, Portugal, 1-3 Dezembro, 2011 (**Anexo 1**).

Agradecimentos

A toda a minha família pelo seu apoio incondicional, em especial aos meus pais por todo o apoio, amor e carinho.

Aos meus colegas de mestrado pela sua amizade e apoio.

Ao Professor João Piedade por todo o seu trabalho, apoio, paciência e incentivo durante toda esta jornada.

Aos elementos do Grupo de Virologia da Unidade de Microbiologia Médica do IHMT, Professora Aida Esteves, Professor Ricardo Parreira e Teresa Venenno pela ajuda e simpatia durante este trabalho. Um especial agradecimento à Sandra Videira e Castro por toda a sua ajuda e orientação ao longo desta jornada.

À Doutora Elizabeth Pádua pela generosa cedência de amostras, utilizadas neste trabalho.

Resumo

O maraviroc (MVC) é o único anti-retroviral antagonista do co-receptor CCR5 licenciado e interage com as ansas transmembranares de CCR5, induzindo uma alteração da sua conformação e impedindo a interacção com gp120. O MVC é activo apenas contra estirpes R5 de HIV-1, sendo utilizado em terapia de recurso.

Neste trabalho, foi estudada a diversidade genética da região C2V3C3 do gene *env* de estirpes de HIV-1 de toxicodependentes por via endovenosa da Grande Lisboa, pesquisando-se também a presença de polimorfismos genéticos naturais. Foram utilizadas 52 amostras de plasma e para 35 destas foi amplificado por RT-*nested* PCR um produto de 565 pb. A análise filogenética revelou a seguinte distribuição de genótipos: 23 B (incluindo, provavelmente, 2 CRF14_BG), 8 A, 3 G e 1 F1. Após tradução, e por comparação com a sequência consenso B, verificou-se uma elevada frequência de polimorfismos genéticos, sendo encontradas algumas “assinaturas de aminoácidos” relativas aos subtipos não-B. Realizou-se ainda uma pesquisa de locais de N-glicosilação e a previsão da utilização de co-receptores (abordagem genotípica), com recurso às regras 11/25 e da carga líquida da ansa V3 e aos programas *PSSM* e *geno2pheno*_[coreceptor]. Observou-se uma conservação genérica do número de locais de N-glicosilação e foram identificadas 5 sequências com tropismo X4 ou duplo. Por fim, com base na literatura, realizou-se uma pesquisa de polimorfismos genéticos associados a resistência ao MVC presentes na ansa V3. Foi observado um número elevado destas mutações. A presença dos padrões 11S+26V e 20F+25D+26V, num total de 3 sequências, é relevante, visto estes estarem inequivocamente associados à resistência *in vivo* ao MVC.

Apesar de não estar ainda definido um perfil de resistência para o MVC, a presença das mutações encontradas, em indivíduos sem contacto prévio com o fármaco, trará implicações relevantes na sua gestão clínica, considerando a introdução do MVC na terapia de recurso.

Abstract

Maraviroc (MVC) is the only CCR5 inhibitor licensed today. This drug interacts with the transmembrane helices of CCR5 co-receptor, inducing a conformation change of its extracellular loops and preventing the interaction with gp120. MVC is only active against R5 strains of HIV-1 and is currently used in salvage therapy.

The genetic diversity of the *env* C2V3C3 region of HIV-1 strains from injecting drug users in the Greater Lisbon was studied, along with the presence of natural genetic polymorphisms. 52 plasma samples were used and the amplification by RT-*nested* PCR of a 565 bp-product was possible in 35 of them. The phylogenetic analysis revealed 23 sequences classified as subtype B (probably including 2 CRF14_BG), 8 A, 3 G and 1 F1. After translation, the presence of natural genetic polymorphisms was studied by comparison to a subtype B consensus. A high frequency of genetic polymorphisms was observed and significant “amino acid signatures” were found in association with non-B subtypes. A full characterization of the N-glycosylation sites was also performed and a coreceptor prediction (genotypic approach) was accomplished using the 11/25 and the V3 net charge rules and the programs PSSM and geno2pheno_[coreceptor]. The number of N-glycosylation sites was generically preserved. Five sequences were defined as X4 or dual-tropic. Based on published data, a search for genetic polymorphisms, present in V3 loop, associated to MVC resistance was finally undertaken. Several of such mutations were observed, being particularly interesting the presence of the patterns 11S+26V and 20F+25D+26V, in a total of 3 sequences, since these patterns have unequivocally been associated with MVC resistance *in vivo*.

Although a resistance profile for MVC is not yet defined, the presence of these mutations in MVC-naïve populations may have significant impact in their clinical management in the future, especially considering the introduction of this drug in salvage therapy.

Índice

Elementos Bibliográficos.....	i
Agradecimentos	ii
Resumo	iii
Abstract	iv
Lista de Abreviaturas.....	viii
1. Introdução	1
1.1. Síndrome da Imunodeficiência Adquirida	1
1.2. Origem e Classificação do HIV.....	2
1.3. Organização da Partícula Viral e do Genoma	3
1.4. Replicação Viral.....	5
1.4.1. Tropismo Celular e Receptores do HIV	5
1.4.2. Ciclo Replicativo	6
1.5. Epidemiologia Molecular do HIV.....	10
1.6. Variabilidade Genética	11
1.7. Terapia Anti-Retroviral.....	12
1.7.1. Inibidores de entrada antagonistas de CCR5	13
1.7.1.1. Maraviroc	15
1.7.1.2. Resistência ao Maraviroc	15
1.8. Objectivos.....	17
2. Material e Métodos.....	19
2.1. Amostras.....	19
2.1.1. Processamento das Amostras	19
2.2. Extracção de RNA.....	20
2.3. Amplificação por RT- <i>nested</i> PCR	21
2.3.1. <i>Primers</i>	21
2.3.2. Protocolo de Retrotranscrição e Amplificação	22
2.3.2.1. Retrotranscrição e Primeira Reacção de PCR	22
2.3.2.2. Segunda Reacção de PCR	23
2.4. Observação dos Fragmentos Resultantes da RT- <i>nested</i> PCR	24
2.5. Purificação dos Fragmentos de DNA	25
2.6. Clonagem dos Produtos de PCR.....	26

2.6.1. Ligação de Fragmentos de DNA ao Vector	26
2.6.2. Preparação de Células Competentes e Transformação	27
2.6.3. Selecção de Clones Bacterianos Recombinantes	28
2.7. Extracção de DNA Plasmídico	28
2.8. Observação do DNA Plasmídico após a Extracção	29
2.9. Digestão com EcoRI	29
2.10. Extracção e Purificação de DNA Plasmídico para Sequenciação	30
2.11. Sequenciação Nucleotídica	31
2.12. Análise Bioinformática de Sequências Nucleotídicas de HIV-1	31
2.12.1. Análise de Cromatogramas e Edição de Sequências	31
2.12.2. Pesquisa de Homologia Genética	31
2.12.3. Determinação Automática dos Genótipos Virais	32
2.12.4. Construção de Árvores Filogenéticas	32
2.12.5. Análise de Sequências Divergentes	32
2.13. Análise Bioinformática de Sequências Aminoácídicas de HIV-1	33
2.13.1. Tradução e Alinhamento das Sequências	33
2.13.2. Análise dos Locais de N-glicosilação	33
2.13.3. Previsão de Utilização de Co-Receptores	33
2.13.4. Análise de Polimorfismos	33
2.14. Análise Estatística	34
3. Resultados	35
3.1. Caracterização da Amostra Estudada	35
3.2. Amplificação por RT- <i>nested</i> PCR	36
3.3. Edição das Sequências Nucleotídicas e Pesquisa de Homologia Genética	38
3.4. Análise Filogenética	38
3.5. Análise das Sequências de Aminoácidos	45
3.5.1. Análise de Polimorfismos Genéticos	48
3.5.2. Locais de N-glicosilação	53
3.5.3. Previsão de Utilização de Co-receptores	55
3.5.4. Análise de Polimorfismos Genéticos Associados a Resistência ao Maraviroc	58
4. Discussão e Conclusões	61

4.1. Amplificação do Genoma do HIV-1	61
4.2. Análise das Sequências Nucleotídicas.....	63
4.3. Análise das Sequências de Aminoácidos	68
4.4. Locais de N-glicosilação	69
4.5. Previsão da Utilização de Co-receptor	71
4.6. Pesquisa de Polimorfismos Genéticos Associados a Resistência ao Maraviroc	74
5. Referências Bibliográficas.....	77
6. Anexos.....	97

Lista de Abreviaturas

aa – Aminoácido(s)

ATP – 5'-trifosfato de adenosina (do inglês *adenosine triphosphate*)

AZT – Azidotimidina

BLAST – *Basic Local Alignment Search Tool*

BSA – Albumina sérica bovina (do inglês *bovine serum albumin*)

CA – Proteína da cápside

CAT – Centro de Atendimento a Toxicodependentes das Taipas

CCR2 – Receptor de quimiocinas da classe CC (do inglês *CC chemokine receptor 2*)

CCR5 – Receptor de quimiocinas da classe CC (do inglês *CC chemokine receptor 5*)

CCR8 – Receptor de quimiocinas da classe CC (do inglês *CC chemokine receptor 8*)

CDC – *Centers for Disease Control and Prevention*

cDNA – Ácido desoxirribonucleico complementar (do inglês *complementary deoxyribonucleic acid*)

CRF – Forma recombinante circulante (do inglês *circulating recombinant form*)

CXCR4 – Receptor de quimiocinas da classe CXC (do inglês *CXC chemokine receptor 4*)

DC-SIGN – *Dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin 1*

DMSO – Dimetilsulfóxido

DNA – Ácido desoxirribonucleico (do inglês *deoxyribonucleic acid*)

DNase – Desoxirribonuclease

dNTP – Desoxirribonucleótido(s)

D.O. – Densidade óptica

DTT – Ditionitreitol

EDTA – Etilenodiaminotetracetato

EMA – *European Medicines Agency*

env – Gene codificante das glicoproteínas do invólucro viral (do inglês *envelope*)

EUA – Estados Unidos da América

FDA – *U.S. Food and Drug Administration*

FI – Inibidor de fusão (do inglês *fusion inhibitor*)

gag – Gene codificante das proteínas da estrutura interna do virião (do inglês *group specific antigen*)
gp – Glicoproteína(s)
 HAART – *Highly active antiretroviral therapy*
 HFF – Hospital Prof. Doutor Fernando Fonseca, EPE
 HIV – *Human immunodeficiency virus*
 HLA – *Human leukocyte antigen*
 HR-1 – Repetição heptavalente hidrofóbica N-terminal (do inglês *N-terminal heptad repeat*)
 HR-2 – Repetição heptavalente hidrofóbica C-terminal (do inglês *C-terminal heptad repeat*)
 ICAM – Molécula de adesão intercelular (do inglês *intercellular adhesion molecule*)
 IDT – Instituto da Droga e da Toxicodependência
 INSA – Instituto Nacional de Saúde Doutor Ricardo Jorge
 IPTG – Isopropil- β -D-tiogalactopiranosídeo
 LB – *Lysogeny broth*
 MA – Proteína da matriz
 MEGA – *Molecular Evolutionary Genetics Analysis*
 mg – Miligrama
 min – Minuto(s)
 ml – Mililitro(s)
 mM – Milimolar
 mRNA – Ácido ribonucleico mensageiro (do inglês *messenger ribonucleic acid*)
 NC – Proteína da nucleocápside
 NCBI – *National Center for Biotechnology Information*
nef – Gene acessório do HIV (do inglês *negative factor*)
 ng – Nanograma(s)
 nm – Nanómetro(s)
 NNRTIs – Inibidores não-nucleosídeos da transcriptase reversa (do inglês *non-nucleoside reverse transcriptase inhibitors*)
 NRTIs – Inibidores nucleosídeos da transcriptase reversa (do inglês *nucleoside reverse transcriptase inhibitors*)

NSI – Não indutor de sincícios (do inglês *non syncytium-inducing*)

O.M.S. – Organização Mundial da Saúde

ORF – *Open Reading Frame*

p – Probabilidade estatística

pb – Par(es) de bases

PCR – Reação em cadeia da polimerase (do inglês *polymerase chain reaction*)

PEG – Polietilenoglicol

PIs – Inibidores da protease (do inglês *protease inhibitors*)

pmoles – Picomoles

pol – Gene codificante das proteínas enzimáticas do HIV (do inglês *polymerase*)

p/v – Peso por volume

rev – Gene regulador do HIV (do inglês *regulator of virion gene expression*)

RNA – Ácido ribonucleico (do inglês *ribonucleic acid*)

RNase – Ribonuclease

rpm – Rotações por minuto

RRE – *Rev response element*

RT – Transcriptase reversa (do inglês *reverse transcriptase*)

SDS – Dodecil sulfato de sódio (do inglês *sodium dodecyl sulfate*)

seg – Segundo(s)

SI – Indutor de sincícios (do inglês *syncytium-inducing*)

sida – Síndrome da imunodeficiência adquirida

SIV – *Simian immunodeficiency virus*

SU – Glicoproteína de superfície

TAE – Tampão Tris-Acetato-EDTA

TAR – *Transactivation response element*

tat – Gene regulador do HIV (do inglês *transcriptional transactivator*)

TEG – Tampão Tris-EDTA-glicose

TM – Glicoproteína transmembranar

T_m – Temperatura de desnaturação (do inglês *melting temperature*)

TSS – Solução de transformação e armazenamento (do inglês *Transformation and Storage Solution*)

U/ μ l – Unidade(s) por microlitro

V – Volt(s)

vif – Gene acessório do HIV (do inglês *viral infectivity factor*)

vpr – Gene acessório do HIV (do inglês *viral protein R*)

vpu – Gene acessório do HIV (do inglês *viral protein U*)

v/v – Volume por volume

x g – Força centrífuga (relativa à aceleração gravitacional terrestre)

X-Gal – 5-bromo-4-cloro-3-indolil- β -D-galactopiranósido

λ – Comprimento de onda

μ g – Micrograma(s)

μ l – Microlitro(s)

°C – Grau Celsius

1. Introdução

1.1. Síndrome da Imunodeficiência Adquirida

Em Junho de 1981, foram reportados vários casos de homossexuais masculinos que apresentavam pneumonia por *Pneumocystis jirovecii*, entre outras infecções oportunistas, sarcoma de Kaposi e linfomas não-Hodgkin (19, 20). Foi constatado que uma grande proporção destes doentes apresentava uma linfadenopatia generalizada, com depleção selectiva de linfócitos T CD4+ (22, 71, 108). Em 1982 foram detectados casos similares, do que viria a ser conhecido como síndrome da imunodeficiência adquirida (sida), em utilizadores de drogas injectáveis, imigrantes haitianos, receptores de transfusões de sangue e produtos sanguíneos, especialmente de factores de coagulação para o tratamento da hemofilia (21, 23, 125). A acompanhar estes casos foram depois identificados casos de transmissão heterossexual (36) e de mãe para filho (24). Estes dados epidemiológicos levaram à conclusão de que a sida poderia ser causada por um agente infeccioso, provavelmente um vírus transmitido por via sexual e sanguínea, uma vez que era pouco provável a contaminação de factores de coagulação por outros tipos de microrganismos (166). O agente etiológico da sida, o vírus da imunodeficiência humana (HIV, do inglês *human immunodeficiency virus*) foi isolado pela primeira vez em 1983 (10).

Actualmente, à entrada da quarta década da pandemia da sida, assiste-se finalmente à reversão do alastramento do HIV. O número de novas infecções tem vindo a diminuir desde 1997, ano em que este valor atingiu o seu máximo, apresentando hoje uma taxa de diminuição de 21% em relação ao valor desse ano (158). A acompanhar esta diminuição no número de novas infecções, resultado do sucesso parcial de acções de educação e prevenção, está o declínio, em cerca de 19%, da taxa de mortalidade de doentes infectados, mesmo em países em desenvolvimento, demonstrando que o acesso alargado à terapia anti-retroviral está a apresentar resultados (158).

Em Portugal, entre todos os casos notificados, para todos os estádios de infecção, os indivíduos que indicam o consumo de drogas por via endovenosa (“toxicoddependência”) como provável causa de infecção constituem 39,9% de todas as notificações, alterando-se em 2010 a tendência inicial da epidemia, uma vez que o número associado à transmissão sexual (heterossexual) representa já 42,1% dos registos acumulados. Os casos notificados de infecção pelo HIV que referem como causa

provável de infecção a transmissão sexual (heterossexual) apresentam uma tendência evolutiva crescente, como se pode avaliar pelos dados disponíveis mais recentes. Em 2010, foram diagnosticados 1020 novos casos de infecção, entre os quais se encontravam 612 heterossexuais, 147 toxicodependentes e 217 homo/bissexuais. Em síntese, nesse ano, as notificações reportam 60% dos novos casos diagnosticados para a categoria de transmissão “heterossexual” e somente 14,4% para a transmissão associada à toxicodependência (81). Estes valores ilustram uma alteração evidente do padrão epidemiológico da sida em Portugal.

1.2. Origem e Classificação do HIV

Actualmente, reconhece-se que os vírus da imunodeficiência humana (HIV-1 e HIV-2) entraram na população humana através de infecções de carácter zoonótico de vírus da imunodeficiência símia (SIV, do inglês *simian immunodeficiency virus*) com origem em primatas infectados (74). Os reservatórios símios dos vírus humanos (HIV-1 e HIV-2) foram definidos com base na similaridade genética e filogenética em relação a SIVs conhecidos e na elevada coincidência entre a distribuição natural de hospedeiros primatas infectados e as regiões geográficas com a maior diversidade genética de HIV-1 e HIV-2 (92), entre outros factores.

O HIV-1 está filogeneticamente mais próximo do SIVcpz_{Ptt} isolado da subespécie de chimpanzés *Pan troglodytes troglodytes* na África Centro-Occidental, (33, 63) local onde é encontrada a maior diversidade genética de HIV-1, enquanto o HIV-2 está evolutivamente mais relacionado com o SIVsm dos mangabeis fuliginosos (*Cercocebus atys*) da África Occidental, sendo que a área de distribuição natural desta espécie é sobreponível com a região geográfica onde foram identificados HIV-2s mais diversos (25, 61).

Os HIV-1 e HIV-2 são membros do género *Lentivirus*, subfamília *Orthoretrovirinae*, da família *Retroviridae*, caracterizando-se genericamente por um passo de transcrição reversa no seu ciclo replicativo. Estes vírus têm exibido excepcionais propriedades de indução de doença insidiosa, persistência, variação genética, recombinação e escape a pressões imunitárias e exercidas por fármacos (53, 74).

1.3. Organização da Partícula Viral e do Genoma

As partículas virais maduras do HIV têm uma morfologia aproximadamente esférica com um diâmetro de cerca de 100-120 nm e possuem um invólucro fosfolipídico (53, 142), associado a uma matriz proteica que circunda, por sua vez, uma cápside com a forma de cone truncado (142). Esta encerra, entre outras, duas moléculas de natureza ribonucleoproteica, resultantes da associação do genoma viral de RNA com as proteínas da nucleocápside (**Figura 1**).

O invólucro viral inclui complexos proteicos triméricos compostos por heterodímeros das glicoproteínas gp120 (SU) e gp41 (TM) ligadas entre si (53). Durante o processo de saída da célula infectada o vírus pode ainda integrar no seu invólucro diferentes proteínas da membrana do hospedeiro, como proteínas HLA classes I e II (68) e proteínas de adesão, como ICAM-1 (52), que podem facilitar a adsorção a outras células alvo (17, 56). A proteína da matriz (p17 ou MA) forma uma estrutura com simetria icosaédrica e está ancorada ao lado interno do invólucro viral. O invólucro viral e a matriz rodeiam a cápside viral composta por monómeros do antígeno p24 (ou CA). Esta contém, por sua vez, o genoma viral, que é composto por duas moléculas idênticas de RNA de cerca de 9,5 kb, de cadeia simples e polaridade positiva, associadas à proteína da nucleocápside (p7 ou NC). No interior da cápside encontram-se ainda as enzimas transcriptase reversa, protease e integrase, as proteínas virais acessórias Vpu, Vif, Vpr, Nef e alguns factores proteicos celulares (58, 142).

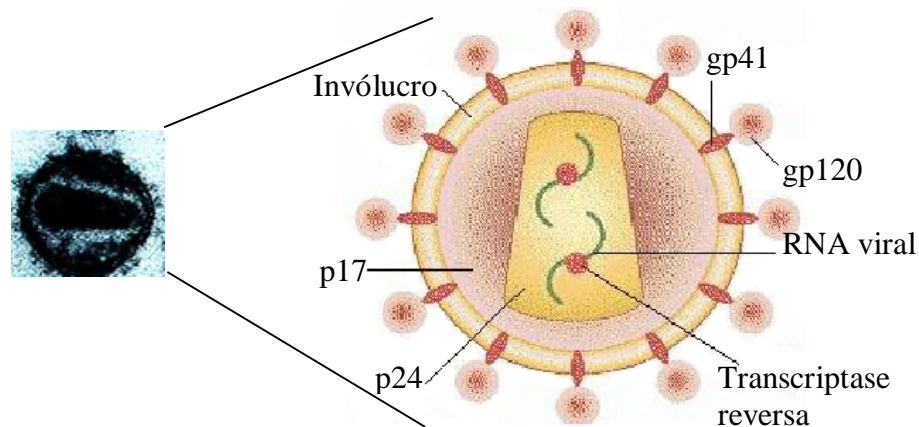


Figura 1: Imagem de microscopia electrónica de uma partícula madura de HIV (<https://qspace.library.queensu.ca/html/1974/136/hiv01.htm>) e respectiva representação esquemática (adaptado de: <http://virology-online.com/viruses/HIV.htm>).

O genoma do HIV-1 é um genoma complexo que possui os genes estruturais *gag*, *pol* e *env* (127) (**Figura 2**). O gene *gag* codifica a poliproteína precursora p55, a qual é clivada pela protease viral nas proteínas maduras da matriz (p17), da cápside (p24), da nucleocápside (p7) e proteína p6, essencial na montagem e liberação dos viriões. As proteínas codificadas pelo gene *pol* são inicialmente sintetizadas como parte de uma poliproteína precursora (p160), também clivada pela protease viral para originar as enzimas virais, protease, transcriptase reversa e integrase. As glicoproteínas do invólucro viral são também sintetizadas como parte de uma poliproteína precursora, gp160, codificada por *env*. Esta é clivada por uma protease celular, resultando na formação das glicoproteínas de superfície (gp120) e transmembranar (gp41) (53, 59), as quais formam um heterodímero. A gp120 compreende os determinantes que interagem com o receptor e co-receptor celulares (59). A gp41 é responsável pelo ancoramento do complexo glicoproteico viral à membrana da célula hospedeira e contém os domínios críticos (domínio fusogénico) para a fusão entre o invólucro viral e a membrana celular, aquando da entrada do vírus.

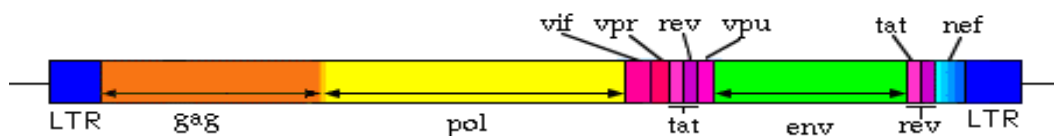


Figura 2: Representação esquemática do genoma do HIV-1 (adaptado de <http://www.yale.edu/bio243/HIV/genome.html>).

A glicoproteína gp120 é composta por um domínio interno, um domínio externo, ligados entre si, e um terceiro elemento designado por “bridging sheet” (93). Nestes domínios estão localizadas cinco regiões conservadas (C1-C5) e cinco regiões variáveis (V1-V5) (113, 170) (**Figura 3**). As regiões variáveis correspondem a ansas expostas na superfície externa da partícula viral (113). Destas regiões, a região V3 é determinante para o tropismo viral (16) e capacidade de fusão do invólucro com a membrana da célula hospedeira (50). A presença de determinados resíduos aminoácidos em posições fixas, neste domínio, determina o co-receptor utilizado pelo vírus e assim o fenótipo apresentado (37, 57, 141). O processo de fusão depende da região V3, mas também das regiões V1, V2 e C4 (12).



Figura 3: Representação esquemática das regiões conservadas (C1-C5) e variáveis (V1-V5) da gp120 do HIV-1 [adaptado de Briz *et al.*, 2006 (15)].

Além dos produtos dos genes *gag*, *pol* e *env*, o HIV-1 codifica também duas proteínas reguladoras [Tat (6) e Rev (54, 145)] e várias acessórias [Nef (4), Vif (88), Vpr (60, 172) e Vpu (29)], com funções diversificadas, mas relacionadas com o controlo do ciclo replicativo viral (53, 58, 59). Na sua forma proviral, o genoma do HIV encontra-se flanqueado por regiões terminais longas não codificantes (LTR, do inglês *long terminal repeats*) contendo diversos promotores transcricionais (**Figura 2**).

1.4. Replicação Viral

1.4.1. Tropismo Celular e Receptores do HIV

Os principais alvos celulares do HIV são os linfócitos T CD4+, os macrófagos e as células dendríticas. Este tropismo é predominantemente determinado pelos receptores celulares de superfície necessários para a ligação e entrada do vírus (27). A identificação, em 1984, da molécula CD4 como receptor utilizado pelo HIV para a entrada nas células (34) foi suportada em várias evidências: 1) a infecção era inibida por anticorpos monoclonais anti-CD4; 2) a expressão de CD4 em células resistentes conferia susceptibilidade à infecção e 3) o receptor CD4 podia ser co-precipitado com o invólucro viral (77).

A molécula CD4 funciona normalmente como co-receptor do complexo maior de histocompatibilidade de classe II durante o reconhecimento pelo receptor das células T de um determinante antigénico estranho (28, 53). CD4 é expresso em cerca de 60% dos linfócitos T circulantes, seus precursores na medula óssea e timo, em monócitos/macrófagos, eosinófilos e células dendríticas (células de Langerhans e células microgliais do sistema nervoso central) (53). Enquanto alguns destes tipos de células apresentam elevados níveis de expressão de CD4 (linfócitos T), outros tipos de células, incluindo macrófagos (28, 104) e células dendríticas, expressam quantidades quase indetectáveis. Neste caso, o HIV pode ligar-se às células através de interações independentes de CD4, utilizando, por exemplo, domínios do tipo lectina em receptores na superfície celular (98).

Após a descoberta de que CD4 é o principal receptor do HIV, surgiram provas de que este não era suficiente para a entrada do HIV nas células (153), ao mesmo tempo que não se encontrava explicado o tropismo variável para diferentes tipos de células CD4+ (59, 165). Foram identificados isolados virais com tropismo preferencial para macrófagos primários e isolados com tropismo preferencial para linfócitos T activados (66), designados também, respectivamente, como NSI (do inglês *non syncytium-inducing*) e SI (do inglês *syncytium-inducing*) (152). Este tropismo diferencial foi explicado com a descoberta de dois co-receptores para a entrada do HIV, os receptores de quimiocinas CCR5 e CXCR4, pertencentes à superfamília de receptores acoplados à proteína G (43, 55). As estirpes com tropismo preferencial para macrófagos utilizam o co-receptor CCR5 e as estirpes com tropismo preferencial para linfócitos T utilizam o co-receptor CXCR4 (3, 55). Assim, com base no co-receptor utilizado, foi desenvolvida uma nova classificação do HIV (11): estirpes que utilizam o co-receptor CCR5 são designadas por estirpes R5, as que utilizam o co-receptor CXCR4 designam-se por X4 e as que conseguem utilizar ambos designam-se por estirpes com tropismo duplo ou R5/X4. Das populações virais que apresentam estirpes R5 e X4, em simultâneo, diz-se que têm tropismo misto. Finalmente, de realçar a existência de estirpes de HIV que utilizam co-receptores alternativos a CCR5 e CXCR4.

Existem também receptores celulares que embora não suportem a fusão do invólucro do HIV com a membrana celular, impossibilitando assim a infecção por essa via, mantêm os viriões à superfície da célula. É o caso do receptor DC-SIGN, presente nas células dendríticas. Após o contacto com o vírus, geralmente ao nível das mucosas ou da pele, e sua ligação, através de gp120, ao receptor DC-SIGN, estas células migram para os gânglios linfáticos regionais onde permitem a *trans*-infecção dos linfócitos T CD4+ (67). No entanto, as células dendríticas, quando no seu estado imaturo, apresentam na sua superfície o receptor CD4 e o co-receptor CCR5, tornando-se assim também susceptíveis à infecção, o que contribui para a proliferação do vírus e infecção de novas células (147).

1.4.2. Ciclo Replicativo

O ciclo replicativo do HIV-1 é genericamente semelhante ao de outros retrovírus, mas existem vários aspectos únicos (59).

O complexo glicoproteico trimérico do envólucro viral, composto de heterodímeros de gp120 e gp41, é essencial para o reconhecimento e entrada na célula alvo. A gp120 liga-se à glicoproteína CD4 (53, 142), após o que sofre modificações estruturais que conduzem à exposição de um domínio específico capaz de se ligar a receptores de quimiocinas na membrana celular. A dupla ligação da gp120 a CD4 e ao co-receptor permite uma ligação mais estável, o que vai permitir que um péptido fusogénico hidrofóbico, localizado na extremidade amino terminal da gp41, penetre na membrana da célula alvo (12, 53, 59). As sequências repetidas HR-1 e HR-2 da gp41 vão interagir entre si causando a alteração conformacional da porção extracelular da gp41 numa formação em forma de gancho de cabelo. Esta estrutura aproxima o envólucro viral e a membrana celular permitindo a fusão das duas membranas (**Figura 4**) e subsequente entrada da cápside viral na célula (53, 59).

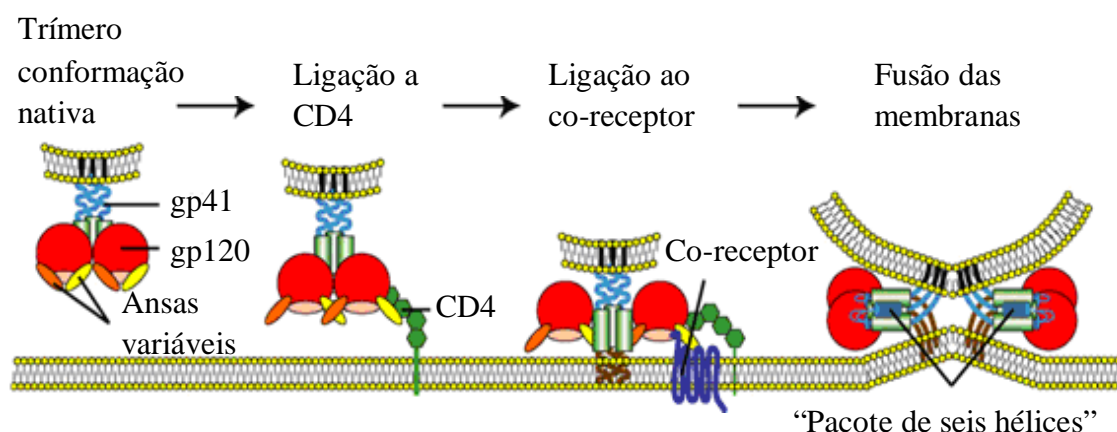


Figura 4: Representação esquemática das fases iniciais da entrada do HIV-1 na célula hospedeira [adaptado de Tilton e Doms, 2010 (155)].

Em seguida, o vírus sofre uma descapsidação parcial no citoplasma da célula, libertando os complexos ribonucleoproteicos do genoma viral (53, 142), dando início à fase da transcrição reversa (**Figura 5**).

A conversão do RNA viral no DNA proviral ocorre por acção da transcriptase reversa e, em fases posteriores do ciclo replicativo, da integrase. A transcriptase reversa, constituída por duas subunidades, catalisa três reacções distintas: a polimerização de uma cadeia simples de DNA, a partir do RNA viral, e para a qual necessita de um iniciador, a degradação concomitante da cadeia de RNA matriz, à excepção de uma pequena região híbrida RNA/DNA, e, finalmente, a polimerização de uma segunda

cadeia de DNA, de modo a formar uma cadeia dupla, utilizando como iniciador a pequena região híbrida mencionada anteriormente. Esta molécula de DNA de cadeia dupla, em associação com várias proteínas virais e celulares, forma o chamado complexo de pré-integração, o qual é transportado para o núcleo da célula infectada. Uma vez no núcleo, este DNA complementar é integrado no genoma da célula pela integrase viral, com intervenção de proteínas do hospedeiro. A formação do DNA proviral e a sua expressão posterior requer que a célula alvo se encontre num estado activado. Monócitos/macrófagos, células microgliais e células T CD4+ quiescentes com infecção latente, contendo o provírus integrado, são reservatórios celulares importantes do HIV.

Aquando da activação celular, inicia-se a transcrição dos RNAs mensageiros de origem viral (53). O processo de transcrição conduz, inicialmente, à síntese das proteínas reguladoras Tat e Rev e da proteína acessória Nef, a partir de mRNAs completamente processados por mecanismos de *splicing* (numa fase inicial só são transcritos estes mRNAs). Este facto permite a acumulação das proteínas reguladoras do ciclo replicativo viral. No núcleo, a Tat liga-se ao local TAR (do inglês *transactivation response element*), presente na extremidade 5' de todos os mRNAs virais, e promove, de um modo indirecto, a transcrição de RNAs virais mais longos (mensageiros e genómicos) e a posterior expressão dos genes estruturais e enzimáticos. Quando a proteína Rev atinge a quantidade necessária, liga-se aos mRNAs não processados através de uma estrutura designada por RRE (do inglês *Rev responsive element*), impedindo o seu processamento total (59, 142). Estes RNAs que não sofrem processamento de *splicing* (mensageiros ou genómicos) ou que sofrem processamento parcial (mensageiros) são depois transportados para o citoplasma, com intervenção da própria proteína Rev. Inicia-se assim a tradução das proteínas virais de expressão mais tardia, codificadas pelos genes *gag*, *pol*, *env*, *vif*, *vpr* e *vpu*.

As proteínas codificadas pelo gene *gag* constituem basicamente as proteínas estruturais do virião, nomeadamente, proteínas da matriz, cápside e nucleocápside, as quais são produzidas a partir de duas poliproteínas precursoras, Pr55^{Gag} e Pr160^{Gag-Pol}, por acção da protease viral. O gene *pol* codifica, a partir da poliproteína precursora Pr160^{Gag-Pol}, as três actividades enzimáticas virais, protease, transcriptase reversa e integrase. O produto do gene *env* (a proteína precursora gp160), após processamento

por uma protease do hospedeiro, dá origem às duas glicoproteínas do invólucro viral. A clivagem das moléculas dos precursores estruturais pela protease viral é necessária para a geração de partículas virais infecciosas.

A formação de novas partículas virais é um processo que ocorre em múltiplos passos, coordenado pelas proteínas codificadas pelo gene *gag* e também por factores celulares. Partículas virais imaturas sofrem gemulação através da membrana da célula hospedeira adquirindo assim o invólucro viral. Durante o processo de gemulação, a membrana lipídica do invólucro viral pode incorporar várias proteínas da célula hospedeira e ser enriquecida com fosfolípidos e colesterol (53, 173).

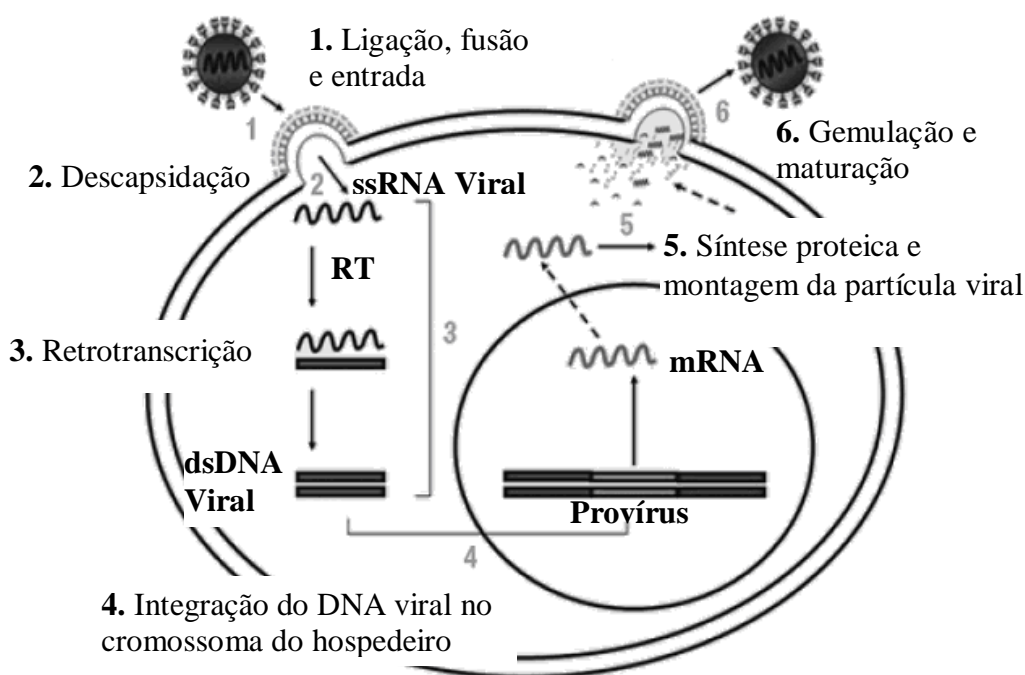


Figura 5: Representação esquemática simplificada do ciclo replicativo do HIV-1 [adaptado de Fanales-Belasio *et al.*, 2010 (53)].

O processo de gemulação em monócitos e macrófagos ocorre de forma diferente do caso dos linfócitos T. Nos monócitos e macrófagos, a gemulação resulta da acumulação de partículas virais em vacúolos intracelulares, de onde são depois libertados. Nos linfócitos T, a gemulação ocorre à superfície da célula e as partículas virais imaturas são libertados para o espaço extracelular (53, 70). A gemulação desencadeia a activação da protease viral, a qual, através de uma afinidade diferencial para determinadas sequências aminoacídicas, inicia uma cascata ordenada de eventos proteolíticos, resultando na clivagem dos precursores Pr55^{Gag} e $\text{Pr160}^{\text{Gag-Pol}}$. Diferentes

interacções que se estabelecem entre as proteínas maduras originam a formação da cápside em forma de cone truncado, característica do HIV, e todas as outras estruturas virais (58, 59, 142).

1.5. Epidemiologia Molecular do HIV

A classificação mais genérica do HIV é em dois tipos, HIV-1 e HIV-2, com origem, como mencionado antes, em dois SIV distintos (SIVcpz_{Ptt} e SIVsm, respectivamente). Cada tipo é classificado em diferentes grupos, correspondentes a diferentes eventos de transmissão inter-específica. Existem três grandes grupos de HIV-1, designados como M (do inglês *major/main*), O (do inglês *outlier*) e N (do inglês *non-M, non-O/new*) (38, 72, 144). Estes grupos representam transferências separadas do vírus de chimpanzés ou, no caso do grupo O, possivelmente de gorilas para humanos (160). Mais recentemente, foi sugerida a existência de um eventual HIV-1 do grupo P (121, 159). O grupo M inclui 9 subtipos, designados como A a D, F a H, J e K (83, 132, 157), alguns dos quais apresentam uma distribuição geográfica relativamente circunscrita na pandemia mundial. Entre os subtipos do grupo M, as distâncias genéticas no gene *env* chegam aos 25-35%, enquanto no gene *gag* podem atingir cerca de 15% (89). Devido à elevada diversidade genética, nos subtipos A e F, podem ainda ser identificados pelo menos 6 sub-subtipos [A1 a A4 (64, 110, 162) e F1 e F2 (156, 157)]. Alguns autores sugerem a criação dos sub-subtipos adicionais A5 e A6, devido à existência de linhagens monofiléticas muito divergentes dentro do subtipo A (8, 163).

Em termos de distribuição geográfica, o subtipo C, o mais frequente à escala mundial, é o mais representado na África Austral e Oriental (36), Índia e algumas regiões da China. O subtipo B é mais prevalente na América, Europa Central e Ocidental e Austrália. O subtipo D está difundido, juntamente com o subtipo A, na África Subsariana Oriental (Uganda, Tanzânia e Quênia). O subtipo A está também presente na Europa Oriental e Ásia, principalmente em crianças e toxicodependentes, respectivamente (53). Em alguns subtipos existem subgrupos, bem localizados geograficamente, com formas virais que partilham o mesmo ancestral, como sugerido em árvores filogenéticas: subtipo B da Tailândia, subtipo C da Índia e Etiópia, subtipo F da Roménia e subtipo G de Espanha e Portugal (49, 89, 154). Os restantes subtipos (H, J, K) encontram-se mais representados, embora em menor frequência, na África Central,

local com a maior diversidade de formas genéticas do vírus. Os grupos O e N, em contraste, estão maioritariamente confinados ao Gabão, Camarões e países vizinhos, nas proximidades do habitat natural de *P.t. troglodytes*, sendo esta região a única onde se encontram em co-circulação todos os grupos (M, N e O) de HIV-1 (8, 126).

Ocasionalmente, dois vírus de diferentes subtipos podem infectar uma mesma célula, constituindo este facto o ponto de partida para a geração de vírus com genomas de tipo mosaico. Muitas destas novas estirpes não têm capacidade replicativa, mas as que são transmitidas, propagando-se com proporções epidémicas, são designadas de CRFs (do inglês *circulating recombinant forms*) (53). Presentemente, estão descritas cerca de 50 CRFs diferentes, sendo as mais prevalentes a CRF01_AE, descoberta inicialmente na Tailândia, e a CRF02_AG, na África Ocidental (13). Uma outra variedade de vírus recombinantes, designada por URFs (do inglês *unique recombinant forms*), corresponde a vírus com genoma mosaico que não se distribuíram epidemicamente, estando confinados a um único indivíduo ou a um grupo de indivíduos epidemiologicamente relacionados (154).

Em Portugal predominam os subtipos G e B e as formas recombinantes circulantes CRF02_AG e CRF14_BG (9, 48, 119).

Relativamente ao HIV-2, este compreende oito grupos, A a H (26, 35, 62, 174), muitos representados apenas por isolados virais únicos. O HIV-2 é endémico na África Ocidental, mas disseminou-se, embora em baixa escala, para a Índia e Europa, principalmente para Portugal e França (53, 89). Foi descrita recentemente a primeira forma recombinante circulante deste vírus, a HIV-2-CRF01_AB (80).

1.6. Variabilidade Genética

A variabilidade genética é de grande importância para o HIV, uma vez que permite ao vírus a fuga à resposta imunitária do hospedeiro e ao efeito de intervenções terapêuticas. Esta elevada variabilidade faz do HIV um dos vírus com a mais elevada taxa de evolução conhecida e esta rápida evolução deve-se a um conjunto de factores: 1) o mecanismo de transcrição propenso a erros da transcriptase reversa, que introduz, em média, uma substituição nucleotídica por genoma por cada ciclo de replicação (1, 106), dado que não tem capacidade de *proofreading* (131); 2) a rápida taxa de replicação viral, que gera um elevado número de viriões por dia (estimado em cerca de 10^{10}) no

indivíduo infectado (120); 3) a ocorrência de recombinação entre dois ou mais genomas virais diferentes presentes no indivíduo infectado, surgindo quando a transcriptase reversa troca de cadeia de RNA durante a retrotranscrição, originando, em média, três eventos de recombinação por genoma por ciclo de replicação (87, 176) e 4) a existência de hipermutação G-A, baseada na actividade da enzima celular APOBEC3 (86, 134). A taxa de evolução é ainda elevada por mecanismos de selecção natural. Assim, amostras recolhidas de um único indivíduo exibem quase sempre uma heterologia significativa (133, 171). Apesar das modificações nucleotídicas estarem distribuídas por todo o genoma, a maior variabilidade ocorre no gene *env*, quer intra quer interindivíduos (73). O termo *quasispecies* é utilizado para descrever o conjunto das populações virais geneticamente heterogéneas presentes num indivíduo infectado por HIV (46).

1.7. Terapia Anti-Retroviral

O tratamento da infecção por HIV teve o seu início com o licenciamento do AZT (azidotimidina) em 1987. O subsequente desenvolvimento dos anti-retrovirais e a definição de programas terapêuticos consistentes e simples constituíram um triunfo para a saúde pública (36). A terapia para o HIV-1 tem progredido significativamente nos últimos anos através da descoberta e desenvolvimento de numerosos inibidores de várias actividades virais. A combinação de três ou mais destes inibidores, em regimes terapêuticos de múltiplos fármacos, designada HAART (do inglês *highly active antiretroviral therapy*), é capaz de inibir quase totalmente a replicação do vírus, atingindo-se valores baixos ou mesmo indetectáveis de HIV-1 em circulação (5, 44).

Actualmente, existem aprovados pela FDA (do inglês *U. S. Food and Drug Administration*) norte-americana 25 anti-retrovirais distintos, apresentados individualmente ou em formas combinadas, que se agrupam em seis categorias principais (<http://www.fda.gov/ForConsumers/byAudience/ForPatientAdvocates/HIVandAIDSActivities/ucm118915.htm>):

1) Inibidores nucleosídeos da transcriptase reversa (NRTIs, do inglês *nucleoside reverse transcriptase inhibitors*), que bloqueiam a actividade da transcriptase reversa ao competirem com os substratos naturais pelo centro activo da enzima, incorporando-se na cadeia nascente de DNA viral complementar (cDNA), actuando como terminadores de cadeia (112);

2) Inibidores não nucleosídeos da transcriptase reversa (NNRTIs, do inglês *non-nucleoside reverse transcriptase inhibitors*), que bloqueiam a actividade de polimerase da enzima através da indução de alterações conformacionais que alteram o seu local catalítico (47, 111);

3) Inibidores da protease (PIs, do inglês *protease inhibitors*), que impedem a clivagem das poliproteínas precursoras do vírus pela protease viral, originando partículas virais estruturalmente desorganizadas e não infecciosas (151);

4) Inibidor de fusão (FI, do inglês *fusion inhibitor*), que se liga à região HR-1 de gp41 impedindo a formação da estrutura “pacote de seis hélices”, necessária para a fusão do invólucro do vírus com a membrana celular (90);

5) Inibidor de entrada – antagonista do co-receptor CCR5, que, uma vez ligado a este co-receptor, impede a sua interacção com a gp120 (44);

6) Inibidor da integrase (II, do inglês *integrase inhibitor*), que interage com o local activo da enzima, impedindo a integração do cDNA viral no cromossoma da célula hospedeira, por inibição da reacção de transferência de cadeia (76).

Apesar da HAART ter transformado radicalmente a abordagem terapêutica da infecção pelo HIV, permanecem ainda muitos problemas. O grande número de comprimidos, dosagens inconvenientes e as toxicidades significativas a longo prazo contribuem para o não cumprimento da terapia (148) e a consequente emergência de vírus resistentes em muitos pacientes (140). Para estes, as opções de tratamento tornam-se limitadas e regimes mais complexos são necessários para impedir a progressão da doença. O aumento da incidência de variantes resistentes, na população infectada por HIV-1 em tratamento, levou ao aumento da transmissão de vírus resistentes a indivíduos que nunca receberam tratamento (175). Um esforço considerável nesta área está agora centrado na descoberta e desenvolvimento de novos inibidores de entrada, disponíveis oralmente. Estes incluem inibidores da ligação da gp120 a CD4 (102) e aos co-receptores CCR5 (44, 96) e CXCR4 (116).

1.7.1. Inibidores de entrada antagonistas de CCR5

O receptor CCR5 apresenta-se como um alvo interessante para intervenção farmacológica, uma vez que indivíduos homozigóticos para uma deleção de 32 pares de

bases no seu gene codificante são resistentes a estirpes R5 de HIV-1 e não apresentam problemas resultantes desta condição, fazendo prever que antagonistas de CCR5 seriam bem tolerados e eficazes no bloqueio da sua interacção com gp120 (103, 136).

Os agentes farmacológicos que bloqueiam a ligação de gp120 ao CCR5 representam uma classe distinta de anti-retrovirais. De facto, estes agentes não são anti-retrovirais no sentido estrito, uma vez que o seu alvo não é qualquer molécula de origem viral, mas sim um receptor celular necessário para a entrada do HIV na célula hospedeira (69).

Existem actualmente três tipos de antagonistas de CCR5 em desenvolvimento (32). Os anticorpos monoclonais anti-CCR5, específicos para o seu domínio N-terminal e/ou a segunda ansa extracelular, competem com a gp120 pela ligação ao co-receptor. Entre estes, o anticorpo PRO140 inibe a entrada do HIV-1 sem inibir a ligação de quimiocinas (82), mas já o anticorpo HGS004 inibe quer a ligação a gp120, quer a de quimiocinas (95). Por outro lado, as quimiocinas modificadas no seu terminal amino conduzem ao bloqueio de CCR5 e à sua internalização prolongada, podendo induzir as vias de sinalização de CCR5 (75). Finalmente, um terceiro tipo de inibidores são as pequenas moléculas antagonistas de CCR5 que interagem com este, impedindo a sua ligação a gp120 (32). Até hoje, foram desenvolvidos e levados a ensaios clínicos em humanos cinco antagonistas de CCR5 – maraviroc (44), vicriviroc (149), aplaviroc (117), INCB009471 (30) e TBR652 (107). Contudo, apenas o primeiro está actualmente aprovado pela FDA e pela EMA (do inglês *European Medicines Agency*) para o tratamento da infecção por HIV-1.

Os inibidores de entrada antagonistas de CCR5 ligam-se numa região hidrofóbica das hélices transmembranares daquele receptor e induzem uma alteração conformacional das ansas extracelulares que interagem com o HIV. A alteração de conformação impede o reconhecimento e consequente ligação da ansa V3 da gp120. Os antagonistas de CCR5 actuam como inibidores alostéricos, uma vez que não se ligam no mesmo local que a gp120 (44), mostrando uma acção muito específica. Assim, estirpes de HIV-1 X4 ou com tropismo duplo são naturalmente resistentes a estes anti-retrovirais.

1.7.1.1. Maraviroc

O maraviroc, com o nome comercial Selzentry ou Celsentri, é uma pequena molécula, selectiva, antagonista da interacção entre o receptor CCR5 e a gp120 de HIV-1, utilizado no tratamento de adultos

infectados. É utilizado em associação com outros anti-retrovirais, apenas nos indivíduos que já tenham efectuado tratamento prévio para a infecção pelo HIV-1 e quando este apresenta

tropismo para CCR5 (51). O maraviroc é quimicamente descrito como 4,4-difluoro-N-[(1S)-3-[exo-3-(3-isopropil-5-metil-4H-1,2,4-triazol-4-il)-8-azabicyclo[3.2.1]oct-8-il]-1-fenilpropil]ciclohexanecarboxamida (**Figura 6**). A sua fórmula molecular é $C_{29}H_{41}F_2N_5O$, apresentando uma massa molecular de 513,67 g/mol.

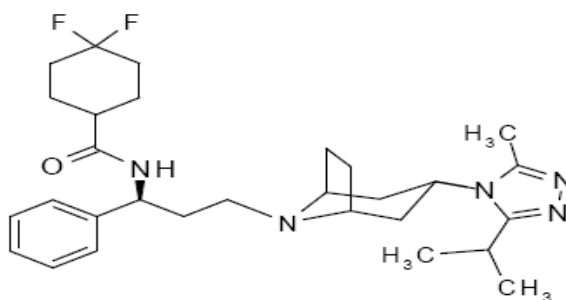


Figura 6: Estrutura química do maraviroc (adaptado de [http://www.viiivhealthcare.com/~media/Files/G/GlaxoSmithKline-Plc/Attachments/pdfs/products/Selzentry-PI-MG-HIV-1](http://www.viiivhealthcare.com/~media/Files/G/GlaxoSmithKline-Plc/Attachments/pdfs/products/Selzentry-PI-MG-HIV-1%2012-13-11.pdf) e [http://www.viiivhealthcare.com/~media/Files/G/GlaxoSmithKline-Plc/Attachments/pdfs/products/Selzentry-PI-MG-HIV-1](http://www.viiivhealthcare.com/~media/Files/G/GlaxoSmithKline-Plc/Attachments/pdfs/products/Selzentry-PI-MG-HIV-1%2012-13-11.pdf) 12-13-11.pdf).

1.7.1.2. Resistência ao Maraviroc

Os mecanismos apontados como possíveis para o desenvolvimento de resistência ao maraviroc e a outros inibidores de CCR5 são: 1) emergência de populações virais X4 pré-existentes; 2) competição com o inibidor pela ligação ao co-receptor através de alterações na gp120 que permitiriam ao vírus utilizar com maior afinidade o co-receptor não ligado ao inibidor (resistência competitiva); 3) utilização do co-receptor ligado ao inibidor, isto é, o vírus adquiriria a capacidade de reconhecer e de se ligar ao co-receptor ligado ao inibidor (resistência não competitiva); 4) utilização de receptores alternativos (CXCR4, CCR2, CCR8, etc.) (128). Entre os diferentes mecanismos possíveis para a resistência ao maraviroc, a alteração de tropismo de R5 para X4, ou para duplo tropismo, e a ocorrência de mutações que possibilitam a utilização do co-receptor ligado ao inibidor, foram as observadas até hoje (167, 168).

A alteração de tropismo é particularmente preocupante uma vez que a utilização do co-receptor CXCR4 está associada a uma rápida diminuição da contagem de linfócitos T CD4⁺ e a uma rápida progressão para sida (31). Esta alteração ocorre

através da emergência de estirpes X4 ou com duplo tropismo, existentes antes do início do tratamento, que passam a ter uma vantagem selectiva em relação a estirpes R5, na presença do inibidor (167). Para que a falência terapêutica por emergência de estirpes X4 pré-existent não ocorra, antes da administração do fármaco deve realizar-se um teste de identificação do co-receptor utilizado pelas estirpes de HIV-1 infectantes. Esta identificação pode ser feita através de um teste fenotípico, disponível comercialmente (Trofile[®] TM, Monogram Biosciences, EUA) (169), ou realizada genotipicamente, baseando-se na sequência nucleotídica do gene *env* (109). Actualmente, no entanto, os testes genotípicos ainda não se encontram disponíveis para uso clínico (69).

No segundo mecanismo observado, a ocorrência de mutações nas regiões codificantes de gp41 e de gp120, especialmente na sua região V3, permite ao vírus a utilização do co-receptor ligado ao inibidor. Este mecanismo é detectado através da observação de um patamar nas curvas de sensibilidade, ou seja, a partir de determinada concentração, o aumento da concentração de inibidor não leva a um aumento da percentagem de inibição (**Figura 7**) (168).

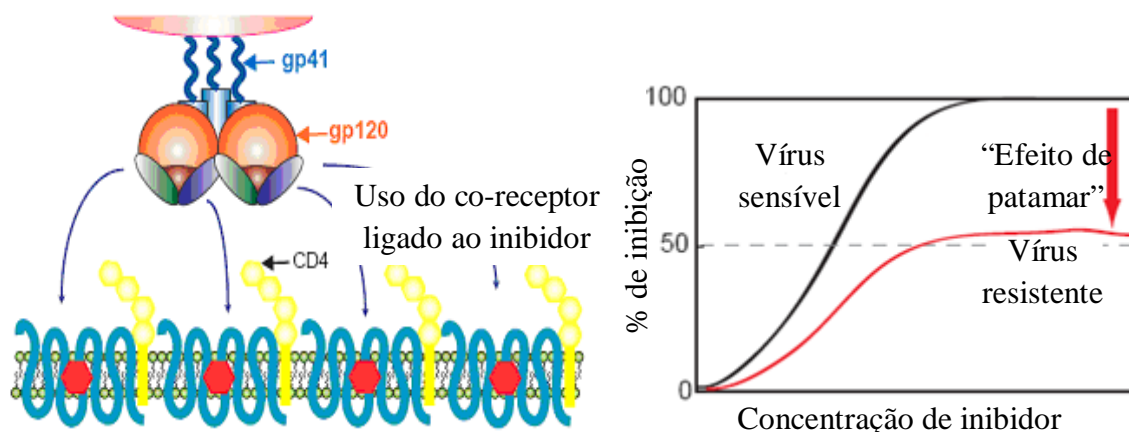


Figura 7: Mecanismo de resistência não competitiva a um inibidor de entrada antagonista de CCR5 e o “efeito de patamar” resultante na curva de dose-efeito [adaptado de Ray, 2008 (128)].

Apesar de terem sido descritas na literatura várias mutações associadas à resistência ao maraviroc (**Tabela 1**), não foi ainda possível definir um padrão de mutações único inequivocamente responsável por esta resistência. Este facto torna problemática a identificação genotípica de resistências a este fármaco (69, 155).

Tabela 1: Mutações associadas à resistência ao maraviroc, compiladas com base na literatura (101, 115, 138, 139).

Numeração HXB2	Mutações ansa V3
297	2I
299	4L
306	11R/S
308	13H/S
313	16A
315	18G
316	19S/T
317	20F/L
318	21I
319	22T
322	25D
323	26V
	G15_P16INSG/A
	N24_I25INSI
	G18DEL
Padrões de Mutações	
	11S+26V
	13H/S+26V
	18G+22T
	19S/T+26V
	20F+25D+26V
	20F+21I
	G15_P16INSG+25D
[Del: Deleção, INS: Inserção]	

1.8. Objectivos

Um dos desafios ao sucesso da terapia anti-retroviral aplicada a indivíduos infectados com o HIV-1 é o aparecimento frequente de estirpes virais naturalmente resistentes aos anti-retrovirais utilizados, mesmo sem qualquer pressão selectiva directa prévia exercida pelo fármaco. O maraviroc, antagonista do co-receptor CCR5, é um dos mais recentes anti-retrovirais licenciados, para o qual os ensaios clínicos realizados estabeleceram que mutações associadas a diminuição significativa da susceptibilidade ocorrem naturalmente em seropositivos nunca tratados. Contudo, a relação entre a presença destas mutações e o estado de resistência clínica *in vivo* mantém-se

controversa, uma vez que há dados contraditórios na literatura e nem sempre os resultados obtidos *in vitro* poderão ser facilmente transpostos para a prática clínica. Por outro lado, os determinantes de resistência ao maraviroc não se encontram totalmente definidos, podendo estar dependentes de outros mecanismos de adaptação do vírus, para além da aquisição de mutações.

Neste âmbito, pretendeu-se com este trabalho:

- 1- Avaliar a diversidade genética das estirpes de HIV-1 circulantes em toxicodependentes por via endovenosa na região da Grande Lisboa, por análise filogenética;
- 2- Analisar a presença de polimorfismos genéticos naturais na região C2V3C3 da gp120 do gene *env*, estimar a sua prevalência e correlacionar com os genótipos virais presentes;
- 3- Estimar a frequência de utilização dos co-receptores CCR5 e CXCR4 (abordagem genotípica), uma vez que a utilização terapêutica do maraviroc estará sempre dependente dessa informação prévia;
- 4- Relacionar os polimorfismos genéticos encontrados na região da ansa V3 com a resistência natural ao maraviroc, por comparação com dados da literatura.

2. Material e Métodos

2.1. Amostras

Neste estudo foi analisado um conjunto de 52 amostras de plasma de indivíduos seropositivos para HIV-1, utilizadores de drogas por via endovenosa. Estas amostras foram obtidas juntamente com o preenchimento de um inquérito que permitiu a compilação de informação de carácter epidemiológico e clínico (**Anexo 2**).

As amostras foram colhidas, entre 1998 e 2009, após obtido o consentimento informado do indivíduo participante no estudo, no:

- Centro de Atendimento a Toxicodependentes das Taipas, Instituto da Droga e da Toxicodependência (CAT das Taipas, IDT)

30 amostras [TP]: PT_TP02, PT_TP03, PT_TP04, PT_TP08, PT_TP13, PT_TP16, PT_TP19, PT_TP24, PT_TP38, PT_TP42, PT_TP44, PT_TP47, PT_TP55, PT_TP63, PT_TP71, PT_TP72, PT_TP85, PT_TP98, PT_TP107, PT_TP114, PT_TP128, PT_TP156, PT_TP158, PT_TP159A, PT_TP184A, PT_TP193, PT_TP208, PT_TP211, PT_TP212, PT_TP219;

- Laboratório de Referência da Sida, Departamento de Doenças Infecciosas, Instituto Nacional de Saúde Doutor Ricardo Jorge (INSA)

21 amostras [TX]: PT_TX12, PT_TX13, PT_TX14, PT_TX15, PT_TX16, PT_TX18, PT_TX20, PT_TX21, PT_TX23, PT_TX30, PT_TX32, PT_TX37, PT_TX40, PT_TX41, PT_TX45, PT_TX46, PT_TX47, PT_TX48, PT_TX50, PT_TX59, PT_TX60;

- Hospital Prof. Doutor Fernando Fonseca, EPE (HFF)

1 amostra [AS]: PT_AS646

As amostras com proveniência no INSA foram gentilmente cedidas pela Doutora Elizabeth Pádua.

2.1.1. Processamento das Amostras

As amostras de 10 ml de sangue periférico, colhidas por punção venosa, foram adicionadas a 0,5 ml de uma solução anticoagulante de citrato de sódio a 3,8% (p/v). No laboratório, foi realizada uma centrifugação de 15 min., a 1800 rpm (*GPR centrifuge*, Beckman, EUA), separando-se assim o plasma. Este (cerca de 5 ml) foi em seguida dividido em alíquotas de 1 ml em tubos de criopreservação. Este procedimento foi

realizado em câmara de fluxo laminar (*Class II Type B3 Bio-Hazard Cabinet*, Enviroco, EUA) num laboratório de nível de segurança 3 (segundo a classificação da O.M.S.). O plasma foi então armazenado a -80°C.

As amostras provenientes do INSA foram fornecidas com a fracção de plasma correspondente, tendo sido armazenadas a -80°C, imediatamente após a sua recepção no laboratório.

Quanto ao estatuto serológico para o HIV-1, este havia sido confirmado anteriormente, por métodos imunoenzimáticos padrão. Por outro lado, estas amostras já tinham sido alvo de estudos de natureza molecular, relativamente a outras regiões do genoma do HIV-1.

2.2. Extracção de RNA

A extracção de RNA viral foi realizada com recurso ao *kit* comercial *High Pure Viral RNA Kit* (Roche Applied Science, Alemanha). Inicialmente, foi preparada a solução de trabalho, por adição de 25 µl de *polyA carrier RNA* (5 mg/ml) a 2,5 ml de tampão de ligação, para cada seis extracções (esta solução deve ser preparada, de fresco, imediatamente antes do início da extracção). Em seguida, colocou-se 400 µl da solução de trabalho em microtubos de 1,5 ml, aos quais foram adicionados 200 µl de plasma, sendo que a manipulação do plasma foi realizada em câmara de fluxo laminar (*Class II Type B3 Bio-Hazard Cabinet*, Enviroco, EUA), em laboratório de nível de segurança 3. Esta mistura foi deixada a incubar durante 10 min. à temperatura ambiente para aumentar o rendimento da extracção. Após a incubação, foi feito um *spin-down* da solução, seguindo-se a sua aplicação a uma coluna (*spin filter* compactado com lã de fibra de vidro) com o respectivo tubo colector. Em seguida foi realizada uma centrifugação de 15 seg. a 8000 x g (*Centrifuge 5415-D*, Eppendorf, Alemanha, utilizada durante todo o processo de extracção). Após a centrifugação, removeu-se a coluna do tubo colector e descartou-se este último juntamente com o eluato, sendo a coluna colocada num novo tubo colector. À coluna adicionou-se 500 µl de tampão de remoção de inibidores, seguindo-se uma centrifugação de 1 min. a 8000 x g. Como anteriormente, o tubo colector foi descartado juntamente com o eluato e a coluna colocada num novo tubo colector. Seguiram-se duas lavagens consecutivas com 450 µl de tampão de lavagem, sendo a coluna centrifugada nas condições anteriores. Após a

segunda lavagem, foi realizada mais uma centrifugação de 10 seg. à velocidade máxima de 13000 x g, para remoção total do tampão de lavagem. A coluna foi finalmente transferida para um novo microtubo de 1,5 ml e foi realizada a eluição do RNA, adicionando-se 50 µl de tampão de eluição, seguindo-se uma centrifugação de 1 min. a 8000 x g. O volume obtido foi em seguida dividido em alíquotas de 10 µl, em tubos livres de RNases e DNases de 0,5 ml (*Multiply®-Pro*, Sarstedt, Alemanha). Esta separação em volumes mais pequenos permite evitar vários ciclos de congelamento e descongelamento do RNA, em futuras utilizações. O passo de eluição foi repetido, tendo-se dividido, desta vez, os 50 µl de eluato em duas alíquotas de 25 µl, sendo este RNA mais diluído que o anteriormente obtido. O RNA obtido foi conservado a -80°C.

Todo o protocolo foi realizado seguindo-se os cuidados necessários ao manuseamento de RNA, que é extremamente susceptível de degradação por RNases endógenas e ambientais.

2.3. Amplificação por RT-nested PCR

2.3.1. Primers

A amplificação da região C2V3C3 de gp120 do gene *env* de HIV-1 foi realizada por transcrição reversa *in vitro* seguida de *nested* PCR (RT-nested PCR). Os *primers* utilizados (**Tabela 2**) encontram-se descritos no protocolo do *HIV-1 env Subtyping Kit* do *NIH AIDS Research & Reference Reagent Program (Division of AIDS, NIAID, NIH, EUA)* para a subtipagem do gene *env* de HIV-1 por *heteroduplex mobility assay* (HMA) (41, 42). Os *primers* seleccionados consistiram em três pares, dois deles utilizados, através de diferentes combinações entre si, para a amplificação de fragmentos maiores que contêm a quase totalidade do gene *env*. A partir destes fragmentos, o terceiro par de *primers*, hibridando nas regiões conservadas (C2, C3) que flanqueiam a região da ansa V3, permite a amplificação de um fragmento final com 565 pb que contém a região V3.

Tabela 2. *Primers* utilizados nas reacções de RT-nested PCR.

	Nome do <i>primer</i>	Posição em HXB2 (pb)	Sequência 5'-3'
<i>Primers</i> Externos (RT e 1 ^a reacção)	ED3	5957–5986	TTAGGCATCTCCTATGGCAGGAAGAAGCG G
	ED5	6557–6582	ATGGGATCAAAGCCTAAAGCCATGTG
	ED12	7811–7782	AGTGCTTCCTGCTGCTCCCAAGAACCCAA G
	ED14	7961–7932	TCTTGCCTGGAGCTGTTTGATGCCCCAGAC
<i>Primers</i> Internos (2 ^a reacção)	ED31	6817–6845	CCTCAGCCATTACACAGGCCTGTCCAAAG
	ED33	7381–7360	TTACAGTAGAAAAATTCCCCTC

2.3.2. Protocolo de Retrotranscrição e Amplificação

2.3.2.1. Retrotranscrição e Primeira Reacção de PCR

Para este estudo recorreu-se ao *Titan One Tube RT-PCR System* (Roche Applied Science, Alemanha). Este sistema permite a realização, num único passo, da reacção de retrotranscrição *in vitro* do RNA numa molécula de cDNA (do inglês *complementary DNA*) e a amplificação deste último. A realização de ambas as reacções, de um modo consecutivo, é possível pela presença de um conjunto de diferentes enzimas, resultando numa amplificação de elevada fiabilidade.

Para a utilização deste sistema foram preparadas duas misturas de reacção, de acordo com as especificações do fabricante e as necessidades deste estudo. Assim, ambas as misturas de reacção foram efectuadas em microtubos livres de RNases e DNases de 0,5 ml (*Multiply[®]-Pro*, Sarstedt, Alemanha), em câmara de fluxo laminar (*Laminar Flow Cabinet-class 100*, Flow Laboratories, EUA), em “sala limpa”. Para a primeira mistura de reacção, colocou-se 1 µl de uma mistura de dNTPs (concentração final, 0,2 mM cada), 2,5 µl de DTT (concentração final, 5 mM), 0,5 µl de RNAGuard™ (Amersham Biosciences, EUA) (concentração final, 0,4 U/µl) e 1,5 µl de cada um dos

primers ED3, ED5, ED12 e ED14 (concentração de 15 pmoles por reacção), acrescentando-se um volume de 5 µl de água livre de RNases (Promega, EUA) para perfazer um volume final de 15 µl. Para a segunda mistura de reacção, juntou-se 10 µl de tampão de RT-PCR 5X (concentração final 1,5 mM de MgCl₂), 1 µl de mistura de enzimas (concentração final de 0,02 U/µl) e 14 µl de água livre de RNases (Promega, EUA) para um volume final de 25 µl. Concluídas as misturas de reacção, e em compartimento físico separado, foram adicionados 10 µl de solução aquosa de RNA purificado à primeira mistura de reacção, sendo depois acrescentados os 25 µl da segunda mistura de reacção, perfazendo o volume final de 50 µl. As misturas de reacção foram mantidas em gelo durante todo o procedimento. Para as reacções de transcrição reversa e amplificação, foi sempre realizado um controlo de contaminação em que se substituiu o volume de RNA/DNA matriz por água.

As condições de amplificação foram escolhidas de acordo com o descrito pelo fabricante do sistema utilizado e tendo em conta as temperaturas de hibridação estimadas dos *primers* utilizados. O protocolo de amplificação foi alvo de várias alterações, ao longo deste trabalho, para sua optimização. As condições descritas na **Tabela 3** foram as que produziram melhores resultados.

Tabela 3. Condições de retrotranscrição e de amplificação (1ª reacção de PCR) do RNA viral.

Retrotranscrição e 1ª reacção de PCR	
Retrotranscrição	42°C, 30 min.
Desnaturação inicial	94°C, 2 min.
Amplificação	94°C, 1 min.; 50°C, 1 min.; 68°C, 1 min. (3X)
	94°C, 10 seg.; 50°C, 30 seg.; 68°C, 1 min., com incremento de 5 seg. por ciclo a partir do 8º ciclo (32X)
Extensão final	68°C, 7 min.
Manutenção	4°C, ∞

2.3.2.2. Segunda Reacção de PCR

Os fragmentos de cDNA resultantes da primeira reacção de amplificação foram utilizados como molde para uma segunda amplificação com *primers* internos (**Tabela**

2). Estes hibridam nos cDNAs permitindo a obtenção de um fragmento de menores dimensões que contém a região em estudo.

Para esta segunda reacção de PCR utilizou-se o sistema *illustra™ puRe Taq Ready-To-Go PCR Beads* (GE Healthcare, Inglaterra) que se apresenta em microtubos de PCR de 0,5 ml contendo uma esfera liofilizada que tem na sua composição aproximadamente 2,5 unidades de polimerase de DNA *puRe Taq*, dNTPs e tampão de reacção. Este método reduz significativamente o número de pipetagens, diminuindo assim a possibilidade de erros e contaminações e aumentando a reprodutibilidade. Foi realizada uma mistura com os dois *primers* internos ED31 e ED33 (15 pmoles/reacção) em água ultrapura *Milli-Q®* (*Reagent Grade Ultrapure Systems*, Millipore, EUA) para um volume de 23 µl por reacção, resultando numa concentração final de 0,2 mM de cada dNTP, 50 mM de KCl e 1,5 mM de MgCl₂ em 10 mM de Tris-HCl, pH 9,0. A esta mistura de reacção foram adicionados 2 µl de solução contendo o cDNA obtido na primeira reacção de amplificação, perfazendo um volume final de 25 µl.

As condições de amplificação utilizadas para esta segunda reacção de PCR foram as descritas no protocolo do *HIV-1 env Subtyping Kit* do *NIH AIDS Research & Reference Reagent Program* (*Division of AIDS*, NIAID, NIH, EUA) e encontram-se na **Tabela 4**.

Tabela 4. Condições de amplificação (2ª reacção de PCR) do cDNA.

2ª reacção de PCR	
Amplificação	94°C, 1 min.; 55°C, 1 min.; 72°C, 1 min. (3X)
	94°C, 15 seg.; 55°C, 45 seg.; 72°C, 1 min. (32X)
Extensão final	72°C, 5 min.
Manutenção	4°C, ∞

2.4. Observação dos Fragmentos Resultantes da RT-nested PCR

A observação dos fragmentos de DNA obtidos por RT-nested PCR foi feita após separação electroforética em gel de agarose (Agarose MP, Boehringer, Alemanha) a 1,5% (p/v), em tampão TAE (Tris-Acetato-EDTA) 0,5X (**Anexo 3**) com brometo de etídeo (0,5 µg/ml), num volume entre 50-100 ml. Foram aplicados 5 µl de amostra com aproximadamente 3 µl de tampão de aplicação (**Anexo 3**), sendo o gel submetido a uma

voltagem a rondar os 90-100 V. O tamanho dos fragmentos obtidos foi estimado por comparação com o padrão de migração do marcador de massa molecular *Gene Ruler™ 100 bp Plus DNA Ladder* (Fermentas, Canadá). A visualização dos fragmentos de DNA foi feita com recurso a radiação ultravioleta ($\lambda=254$ nm) e a captação digital de imagem foi feita pelo aparelho *Gel-Doc XR* (Bio-Rad, EUA).

2.5. Purificação dos Fragmentos de DNA

Após a confirmação da presença de uma banda específica com o tamanho esperado, por comparação com o marcador de massa molecular, e que esta se apresentava com uma quantidade de DNA razoável para sequenciação nucleotídica directa, procedeu-se à purificação do produto de PCR restante com recurso ao sistema comercial *Wizard® SV Gel and PCR Clean-Up System* (Promega, EUA), de acordo com as especificações do fabricante. A 20 μ l de volume de produto de PCR foi adicionado um igual volume de solução de ligação à membrana. Estes 40 μ l foram transferidos para uma minicoluna, já acoplada a um tubo colector, e incubados à temperatura ambiente durante 1 min. Seguiu-se uma centrifugação de 1 min. a 16000 x g, à temperatura ambiente (*Centrifuge 5415-D*, Eppendorf, Alemanha, utilizada durante todo o procedimento de purificação). Retirou-se a minicoluna do tubo colector e rejeitou-se o eluato, retornando-se a minicoluna para o respectivo tubo colector. Em seguida foi realizada uma lavagem com a adição de 700 μ l de solução de lavagem da membrana e centrifugou-se durante 1 min. a 16000 x g à temperatura ambiente. Como anteriormente, após a centrifugação retirou-se a minicoluna do tubo colector, descartou-se o eluato e voltou-se a colocar a minicoluna no respectivo tubo. A lavagem foi repetida com 500 μ l da mesma solução, seguida por uma centrifugação a 16000 x g durante 5 min., rejeitando-se o eluato. O conjunto minicoluna e tubo colector sofreram nova centrifugação por 1 min., nas condições descritas anteriormente, para remover todos os resíduos de solução de lavagem (etanol) que estivessem ainda presentes na coluna. A minicoluna foi em seguida transferida para um microtubo de 1,5 ml, sendo aplicados, no centro de coluna, sem tocar na membrana, 50 μ l de água livre de RNases (Promega, EUA) e deixando-se 1 min. à temperatura ambiente. Foi realizada uma centrifugação final de 1 min. a 16000 x g à temperatura ambiente. A coluna foi em seguida descartada e o DNA purificado conservado a -20°C.

2.6. Clonagem dos Produtos de PCR

Este procedimento foi realizado para todos os produtos de PCR para os quais se observou em gel de agarose possuírem uma quantidade de DNA insuficiente para sequenciação nucleotídica directa. A clonagem foi realizada utilizando o sistema *pGEM[®]-T Easy Vector* (Promega, EUA), cujo plasmídeo se encontra representado na **Figura 8**.

O vector *pGEM[®]-T Easy* é fornecido numa forma linearizada com resíduos de 3'-timidina em ambas as extremidades. A presença destes resíduos aumenta significativamente a eficiência de ligação dos produtos de PCR, ao mesmo tempo que impede a recircularização do vector. O vector permite a selecção de recombinantes através do sistema de α -complementação, uma vez que a inserção de um fragmento no

local de clonagem interrompe a ORF (do inglês *Open Reading Frame*) do gene *lacZ* impedindo assim a produção de β -galactosidase, tornando brancas as colónias bacterianas recombinantes, em meio contendo X-Gal (5-bromo-4-cloro-3-indolil- β -D-galactopiranósido) e IPTG (isopropil- β -D-tiogalactopiranósido) (**Anexo 3**). Este vector permite ainda a fácil extracção do fragmento de DNA clonado com recurso a locais de reconhecimento para enzimas de restrição (EcoRI, BstZI e NotI) que flanqueiam a região de clonagem.

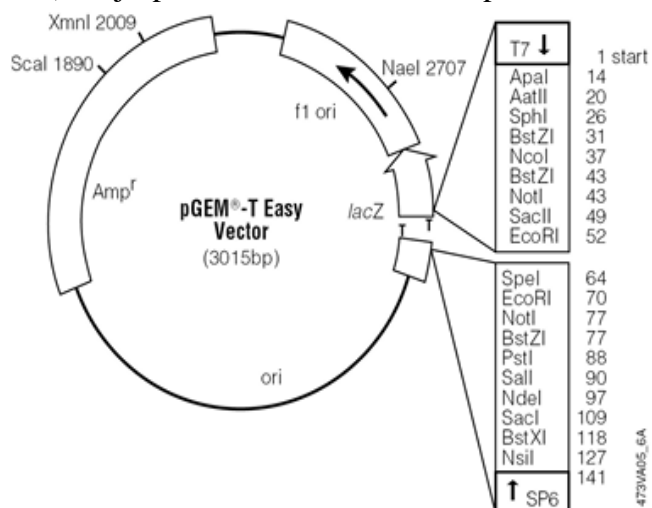


Figura 8: Mapa físico do vector plasmídico *pGEM[®]-T Easy* (adaptado de http://www.promega.com/products/pcr/pcr-cloning/pgem_t-easy-vector-systems/).

2.6.1. Ligação de Fragmentos de DNA ao Vector

A reacção de ligação foi realizada segundo as especificações do fornecedor do vector. Para tal, misturou-se 5 μ l de tampão de ligação 2X (**Anexo 3**), 1 μ l de solução de vector plasmídico *pGEM[®]-T Easy* (Promega, EUA) (50 ng/ μ l), 3 μ l de produto de PCR e 1 μ l de ligase de DNA do fago T4 (3 U/ μ l). Esta mistura foi incubada em banho

refrigerado (*Frigiterm*, P-Selecta, Espanha), a 4°C, durante a noite, de modo a aumentar a eficiência de ligação.

2.6.2. Preparação de Células Competentes e Transformação

Para a obtenção de células competentes para transformação utilizou-se a estirpe JM109 de *Escherichia coli*, previamente conservada a -80°C em meio de criopreservação (**Anexo 3**). Esta estirpe foi inoculada em 3 ml de meio LB (do inglês *lysogeny broth*) líquido (**Anexo 3**) e incubada durante a noite a 37°C, com agitação de 220 rpm. No dia seguinte, utilizaram-se 200 µl da suspensão celular saturada para inocular 20 ml de meio LB fresco num nefalómetro, que se utilizou para a leitura espectrofotométrica da densidade óptica (D.O.) a 600 nm até se obter um valor compreendido entre 0,3 e 0,5, após incubação nas condições anteriores. Uma vez atingida a D.O. pretendida, a cultura foi centrifugada a 4°C, a 6000 rpm durante 10 min. (*Centrifuge J-21*, Beckman, EUA) com o objectivo de sedimentar as células. Em seguida, foi descartado todo o sobrenadante, procedendo-se à ressuspensão das células em 1 ml de meio TSS (do inglês *transformation and storage solution*) frio (**Anexo 3**).

Para a transformação propriamente dita, foram adicionados 5 µl de mistura de ligação (ver 2.6.1) a 100 µl da suspensão de células competentes, em meio TSS, preparada anteriormente. Este procedimento repetiu-se para os controlos positivo e negativo da experiência de transformação, substituindo-se a mistura de ligação respectivamente por 1 µl de uma solução concentrada do plasmídeo *pBluescript II KS+* (Stratagene, EUA) e 5 µl de água ultrapura *Milli-Q®* (*Reagent Grade Ultrapure Systems*, Millipore, EUA). Após homogeneização suave, os microtubos foram mantidos em gelo durante 60 min. Procedeu-se depois a um choque térmico, pelo que as células foram colocadas num banho a 42°C durante 90 seg., sendo imediatamente transferidas para o gelo. Em seguida, adicionou-se 900 µl de meio LB a cada microtubo e colocou-se num agitador orbital (~80 rpm) durante 60 min. a 37°C. No final da incubação, procedeu-se a uma centrifugação de 10 min. a 6000 rpm à temperatura ambiente (*Centrifuge 5415-D*, Eppendorf, Alemanha), sendo cuidadosamente rejeitados 800 µl de sobrenadante e procedendo-se à ressuspensão das células sedimentadas nos 200 µl restantes. Este volume foi finalmente distribuído por espalhamento, em condições de assepsia, em duas placas de Petri com meio LB sólido (15 ml/placa) suplementado com

ampicilina, IPTG e X-Gal (**Anexo 3**). As placas foram colocadas na estufa a 37°C, durante a noite.

2.6.3. Selecção de Clones Bacterianos Recombinantes

A selecção dos transformantes obtidos foi realizada através da análise das placas de Petri incubadas durante a noite. A adição de suplementos ao meio LB sólido permitiu a selecção dos clones transformantes (resistentes à ampicilina) e recombinantes (colónias brancas na presença de X-Gal e IPTG). Entre as colónias obtidas foram seleccionadas colónias brancas para repicagem, em condições de assepsia, com um palito estéril, para tubos com 3,5 ml de meio LB líquido com ampicilina (100 µg/ml). Os tubos foram deixados a incubar durante a noite a 37°C com agitação de 220 rpm.

2.7. Extracção de DNA Plasmídico

Para a extracção do DNA plasmídico dos clones bacterianos seleccionados, utilizou-se o método da lise alcalina (14). Transferiu-se 2x1,5 ml de cultura densa para microtubos de centrífuga e procedeu-se à sua centrifugação a 13500 rpm durante 2 min. (*Centrifuge* 5415-D, Eppendorf, Alemanha), em dois passos consecutivos de centrifugação, com o intuito de sedimentar as células (os 0,5 ml de cultura restantes foram conservados a 4°C). No final, o sedimento celular correspondente a um volume de cultura de 3,0 ml foi vortexado, a fim de facilitar a sua posterior ressuspensão. Em seguida, adicionou-se 100 µl de tampão TEG gelado (**Anexo 3**), procedendo-se à ressuspensão do sedimento celular nesta solução. A esta solução adicionou-se 200 µl de SDS alcalino (**Anexo 3**), preparado no momento, e procedeu-se à mistura através da inversão dos microtubos, de modo a facilitar a lise celular. Finalmente, adicionou-se 200 µl de acetato de potássio 3 M, pH 4,8, frio e inverteu-se os microtubos suavemente para se proceder à mistura. Todo este procedimento foi realizado mantendo os microtubos em gelo. Seguidamente, procedeu-se à sedimentação dos restos celulares, através de uma centrifugação a 13500 rpm durante 15 min., a 4°C (*Centrifuge* 3k18, Sigma, Alemanha). O sobrenadante obtido foi transferido para novos microtubos aos quais se adicionou 600 µl de isopropanol (Merck, Alemanha), procedendo-se à homogeneização por inversão. Esta mistura foi incubada até 10 min. à temperatura ambiente. No final da incubação, realizou-se uma centrifugação a 13500 rpm durante 15

min. à temperatura ambiente. (*Centrifuge 5415-D*, Eppendorf, Alemanha), para recolha do DNA plasmídico. O sobrenadante resultante foi rejeitado e o sedimento de DNA foi lavado com 250 µl de etanol (Merck, Alemanha) a 70% (v/v), mantido a -20°C, realizando-se nova centrifugação, durante 5 min., nas condições anteriores. O sobrenadante foi cuidadosamente retirado e o sedimento de DNA plasmídico foi seco num evaporador (*Concentrator 5301*, Eppendorf, Alemanha). Os sedimentos secos foram, por fim, ressuspensos numa solução aquosa de RNase (concentração final de 50 µg/ml), sendo colocados num banho a 37°C durante 30 min. No final da hidrólise do RNA, o DNA plasmídico foi conservado a -20°C.

2.8. Observação do DNA Plasmídico após a Extracção

Após a extracção do DNA plasmídico, foi realizada uma electroforese em gel de agarose (Agarose MP, Boehringer, Alemanha) a 1% (p/v), em tampão TAE 0,5X, (**Anexo 3**) com brometo de etídeo (0,5 µg/ml), para pré-avaliação do sucesso da clonagem. Os padrões de migração dos DNAs plasmídicos purificados foram comparados, entre si, a fim de se seleccionar as moléculas potencialmente recombinantes. Através desta comparação, seleccionou-se um conjunto de DNAs plasmídicos com migração retardada para digestão posterior com EcoRI, enzima que permite a excisão do fragmento clonado no vector (ver secção 2.6.).

2.9. Digestão com EcoRI

Para a confirmação da presença do fragmento clonado, foi realizada a hidrólise dos plasmídeos apresentando mobilidade retardada em campo eléctrico com a endonuclease de restrição EcoRI (Fermentas, Canadá). Como se viu, esta enzima reconhece dois locais situados no sítio de clonagem múltipla, os quais flanqueiam o local de inserção do fragmento de DNA (**Figura 8**), permitindo a sua excisão. A hidrólise dos DNAs plasmídicos seleccionados foi realizada através da mistura de 5 µl de solução de DNA extraído, 1,5 µl de tampão EcoRI 10X (**Anexo 3**), 1 µl de EcoRI (10 U/µl) e água desionizada ultrapura *Milli-Q®* (*Reagent Grade Ultrapure Systems*, Millipore, EUA) a perfazer um volume total de 15 µl. Esta mistura foi incubada a 37°C durante 3 horas. Os produtos de digestão foram separados por electroforese em gel de agarose (Agarose MP, Boehringer, Alemanha) a 1,5% (p/v), em tampão TAE 0,5X, com

brometo de etídeo (0,5 µg/ml) e analisados após exposição a radiação ultravioleta, nas condições descritas anteriormente.

2.10. Extracção e Purificação de DNA Plasmídico para Sequenciação

Após a visualização de um padrão de restrição contendo fragmentos de DNA com tamanhos aproximados aos teóricos esperados (vector + fragmento clonado), procedeu-se à inoculação de 3,5 ml de meio LB líquido com ampicilina (100 µg/ml), a partir das culturas obtidas por repicagem, e incubação a 37°C durante a noite, com agitação (220 rpm). No dia seguinte, a extracção e purificação do DNA plasmídico para sequenciação nucleotídica foi realizada com recurso ao *kit Wizard® Plus SV Minipreps DNA Purification System* (Promega, EUA), de acordo com as especificações do fabricante. Iniciou-se a extracção com a centrifugação consecutiva de 2x1,5 ml de meio de cultura saturado a 10000 x g durante 5 min. (*Centrifuge 5415-D*, Eppendorf, Alemanha, utilizada durante todo o processo de purificação), para sedimentação das células. Ao sedimento, adicionou-se 250 µl de solução de ressuspensão celular, procedendo-se à ressuspensão por pipetagem repetida. Depois, adicionou-se 250 µl de solução de lise celular, procedendo-se à mistura por inversão e incubação à temperatura ambiente (máximo 5 minutos) até completa clarificação, e 10 µl de uma solução de protease alcalina (~250 µg), também misturada por inversão, sendo os microtubos incubados durante 5 min. à temperatura ambiente. Para a neutralização, adicionou-se 350 µl de solução de neutralização, invertendo-se os microtubos para uma correcta mistura das soluções. Procedeu-se em seguida a uma centrifugação a 14000 x g durante 10 min. à temperatura ambiente. O sobrenadante obtido foi então transferido (cerca de 850 µl) para uma coluna (*spin column*), montada num tubo colector, ambos fornecidos no *kit*, após o que se seguiu uma centrifugação a 14000 x g durante 1 min., à temperatura ambiente. O eluato recolhido no tubo colector foi rejeitado e a coluna voltou a ser colocada no mesmo tubo. À coluna adicionou-se 750 µl de solução de lavagem, procedendo-se a uma centrifugação, nas condições anteriores. O eluato foi descartado, sendo repetida a lavagem da coluna, mas, desta vez, com a adição de 250 µl de solução de lavagem. A coluna foi finalmente transferida para um microtubo estéril de 1,5 ml, procedendo-se à eluição do DNA plasmídico através da adição de 100 µl de água livre de RNases e DNases (Promega, EUA) ao centro da coluna, sem lhe tocar com

a ponta da pipeta. Seguiu-se uma centrifugação a 14000 x g durante 1 min. à temperatura ambiente e o DNA plasmídico foi conservado a -20°C.

2.11. Sequenciação Nucleotídica

Uma vez purificadas, as amostras de DNA foram enviadas para sequenciação nucleotídica na empresa STAB VIDA (Caparica, Portugal). O volume de amostra enviado, quando esta correspondia a produto de PCR, foi variável, dependendo da concentração inicial do DNA (quantidade mínima de 400 ng de DNA/reacção de sequenciação). Os *primers* ED31 e ED33 foram enviados para obtenção de sequências nucleotídicas nos dois sentidos (50 pmoles/reacção de sequenciação). Quanto aos fragmentos clonados, foram enviados 95 µl de solução de DNA plasmídico purificado (quantidade mínima de 1 µg de DNA/reacção de sequenciação). Neste caso, foram utilizados os *primers* universais T7 (directo) e SP6 (inverso), complementares às sequências que flanqueiam o sítio de clonagem múltipla do vector *pGEM[®]-T Easy* (Promega, EUA) (**Figura 8**). Em ambos os casos, a sequenciação foi realizada pelo método de terminação de cadeia (137).

2.12. Análise Bioinformática de Sequências Nucleotídicas de HIV-1

2.12.1. Análise de Cromatogramas e Edição de Sequências

Os cromatogramas relativos às reacções de sequenciação foram analisados a fim de se proceder à edição de sequências “consenso” entre as sequências das cadeias directa e inversa de cada amostra. Empregou-se, para tal, o programa *BioEdit Sequence Alignment Editor*, v.7.0.9.0. (em <http://www.mbio.ncsu.edu/BioEdit/bioedit.html>).

2.12.2. Pesquisa de Homologia Genética

Com o objectivo de se confirmar rapidamente a identidade das sequências obtidas realizou-se uma pesquisa de homologia genética, utilizando-se para tal o programa *BLAST (Basic Local Alignment Search Tool)* (disponível em <http://blast.ncbi.nlm.nih.gov/Blast.cgi>). Este programa indica a percentagem de semelhança entre duas sequências e fornece uma estimativa da probabilidade da homologia genética encontrada ser devida ao acaso (*E value*, 0-1).

2.12.3. Determinação Automática dos Genótipos Virais

Para uma determinação preliminar dos genótipos virais recorreu-se ao programa *REGA HIV-1 & 2 Automated Subtyping Tool v.2.0* (disponível em <http://www.bioafrica.net/rega-genotype/html/subtypinghiv.html>). Este programa realiza uma análise filogenética automática rápida, utilizando métodos de *bootstrapping* e de *bootscanning* para a genotipagem de fragmentos subgenómicos localizados, embora, esta última, apenas para sequências com o tamanho mínimo de 800 pb (39).

2.12.4. Construção de Árvores Filogenéticas

A construção de árvores filogenéticas permitiu a genotipagem das amostras em estudo. Para tal, as sequências nucleotídicas editadas anteriormente foram alinhadas utilizando o programa *ClustalW2 v.2.0*. (disponível em <http://www.ebi.ac.uk/Tools/msa/clustalw2/>). Neste alinhamento foram utilizadas as sequências em estudo, assim como sequências de referência do HIV-1 do grupo M (de subtipos, sub-subtipos e duas CRFs) e uma sequência de SIV, como *outgroup*, obtidas através da *Los Alamos HIV Database* (em <http://www.hiv.lanl.gov/>). O alinhamento obtido foi editado manualmente no programa *BioEdit v.7.0.9.0* para remoção dos espaços abertos. A construção de árvores filogenéticas foi efectuada com o programa *MEGA (Molecular Evolutionary Genetics Analysis) v.4.0* (disponível em <http://www.megasoftware.net/>), com recurso ao método de *neighbor-joining* (135), implementado através do modelo de 2 parâmetros de Kimura (91) para o cálculo da matriz de distâncias genéticas (assumiu-se a opção *pairwise deletion*, uma taxa de transversões igual à taxa de transições e uma frequência igual de bases). Finalmente, foi utilizado o método de *bootstrapping*, com base em 1000 amostragens aleatórias, para avaliar a robustez da inferência filogenética e das topologias sugeridas. Considerou-se como estatisticamente significativo, a obtenção de valores de *bootstrap* iguais ou superiores a 70%.

2.12.5. Análise de Sequências Divergentes

Para a análise de sequências filogeneticamente divergentes, potencialmente recombinantes, utilizou-se o método de *bootscanning*, através da utilização do programa *SimPlot v.3.5.1* (disponível em <http://sray.med.som.jhmi.edu/SCSoftware/simplot/>).

2.13. Análise Bioinformática de Sequências Aminoacídicas de HIV-1

2.13.1. Tradução e Alinhamento das Sequências

As sequências nucleotídicas foram traduzidas segundo o código genético universal com recurso ao programa *EMBOSS Transeq* (disponível em http://www.ebi.ac.uk/Tools/st/emboss_transeq/). Posteriormente, as sequências de aminoácidos foram alinhadas com uma sequência consenso do subtipo B de HIV-1, retirada da base de dados de sequências de HIV (*Los Alamos HIV Database*), utilizando o programa *ClustalW2* v.2.0. e a matriz de substituição BLOSUM 62.

2.13.2. Análise dos Locais de N-glicosilação

Para a análise dos potenciais locais de N-glicosilação (NxT/S, em que N = asparagina, T = treonina, S = serina e x pode representar qualquer aminoácido, excepto prolina) utilizou-se o programa *N-GlycoSite*, disponível em <http://www.hiv.lanl.gov/content/sequence/GLYCOSITE/glycosite.html> (177).

2.13.3. Previsão de Utilização de Co-Receptores

Inicialmente, analisou-se a carga aminoacídica com recurso às regras empíricas 11/25 (57) e carga da ansa V3 (37). A primeira consiste na análise da carga dos resíduos de aminoácidos presentes nas posições +11 e +25 da ansa V3 da gp120 do gene *env*. A segunda regra baseia-se na análise da carga líquida total da mesma região genómica. Em seguida, alargou-se esta análise à restante região V3 utilizando-se os programas *FORTINBRAS PSSM* [matriz: subtipo B X4/R5 (Poveda2009)] (84, 85, 124) (disponível em <http://fortinbras.us/cgi-bin/fssm/fssm.pl>) e *geno2pheno_[coreceptor]* v.1.2. (disponível em <http://coreceptor.bioinf.mpi-inf.mpg.de/index.php>), com um nível de significância de 20%, de acordo com as recomendações (de 2011) do *European Consensus Group* para a gestão clínica dos testes de tropismo de HIV-1, entretanto alteradas (161).

2.13.4. Análise de Polimorfismos

Foi realizada uma pesquisa de polimorfismos naturais, posição a posição, nas sequências em estudo, por comparação com a sequência consenso do HIV-1 do subtipo B utilizada no alinhamento, estimando-se a respectiva prevalência. Por fim, relacionou-

se os polimorfismos encontrados com os descritos na literatura em associação à resistência ao maraviroc (ver **Tabela 1**).

2.14. Análise Estatística

Uma análise de distribuição foi aplicada aos dados epidemiológicos, de natureza categórica, que caracterizam a amostra populacional estudada, como sexo, origem geográfica dos indivíduos, partilha de seringas e terapêutica anti-retroviral. Relativamente à variável numérica “idade dos indivíduos”, calculou-se a média e o desvio padrão. Estudos de associação entre o sucesso/insucesso de amplificação por RT-*nested* PCR e as características epidemiológicas e clínicas da amostra populacional em estudo, e de polimorfismos genéticos, “assinaturas aminoacídicas”, locais de N-glicosilação, mutações associadas a resistência ao maraviroc, entre outras variáveis, com os diferentes subtipos virais (B vs. não-B), ou outras características, foram realizados recorrendo-se ao teste exacto de Fisher bilateral, ou ao teste de χ^2 , aplicados à análise de tabelas de contingência $n \times c$, considerando um valor de significância de 1%. A comparação de médias foi realizada com utilização do teste t de Student. Todos os testes estatísticos foram implementados através de <http://www.physics.csbsju.edu/stats/>.

3. Resultados

3.1. Caracterização da Amostra Estudada

Para este estudo, foi seleccionado um conjunto de amostras de plasma de indivíduos utilizadores de drogas por via endovenosa, seropositivos para HIV-1, da região da Grande Lisboa (ver Material e Métodos). Estes indivíduos apresentavam idades compreendidas entre os 24 e os 53 anos, correspondendo a uma idade média de $37,3 \pm 5,84$ anos. Das 52 amostras utilizadas, 34 pertenciam a indivíduos do sexo masculino e 18 a indivíduos do sexo feminino (**Figura 9**).

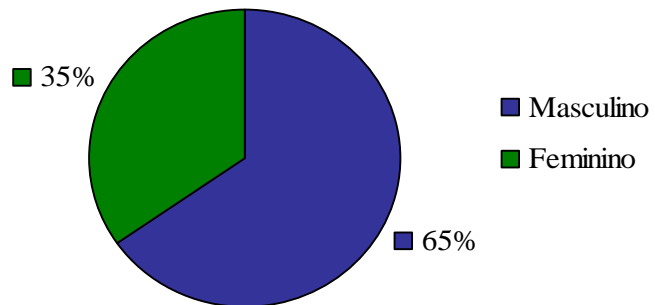


Figura 9: Representação gráfica da distribuição por sexos dos indivíduos cujas amostras foram estudadas.

As amostras estudadas neste trabalho pertenciam maioritariamente a indivíduos de nacionalidade portuguesa (67%), mas englobavam também outras nacionalidades, fundamentalmente de origem africana, como se pode observar na **Figura 10**.

Entre os indivíduos, 65% expressaram afirmativamente a partilha de seringas, representando este aspecto especial importância, dado ser um dos comportamentos de risco da população estudada (**Figura 11**). Para além deste

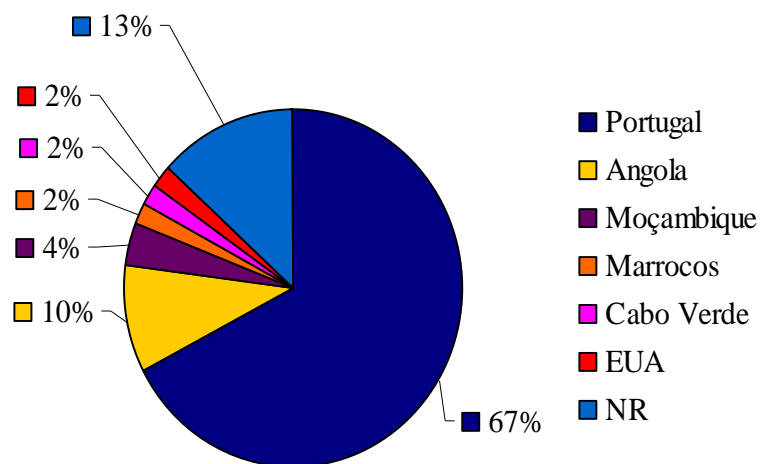


Figura 10: Representação gráfica da origem geográfica dos indivíduos cujas amostras foram estudadas (NR: não reportado).

comportamento, 46% dos indivíduos reportaram a prática de relações sexuais não protegidas. A existência de parceiros sexuais de origem geográfica não-portuguesa foi relatada por 21% da amostra. Apenas 25% dos indivíduos afirmou alguma vez ter

estado sob algum regime terapêutico para o HIV (**Figura 12**). Finalmente, e sendo estes apenas conhecidos para as amostras provenientes do INSA (n=21), os valores de carga viral variavam entre 610 e $1,9 \times 10^6$ cópias de RNA/ml (valor médio de $1,3 \times 10^5 \pm 4,09 \times 10^5$) (**Anexo 2**).

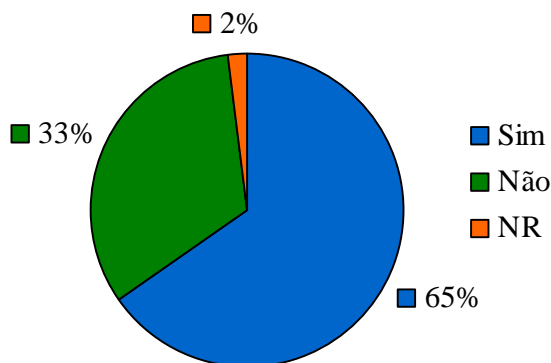


Figura 11: Representação gráfica da prática de partilha de seringas pelos indivíduos cujas amostras foram estudadas (NR: não reportado).

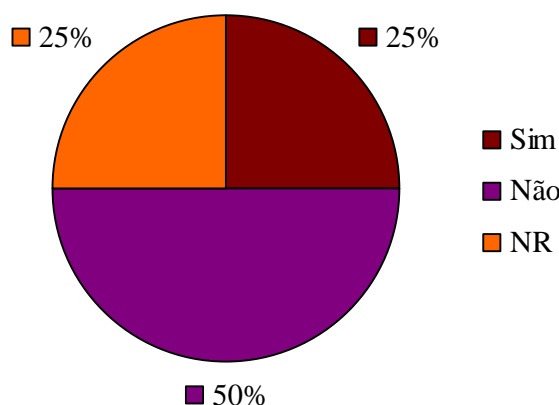


Figura 12: Representação gráfica da percentagem de indivíduos que afirmam ter recebido ou não algum tipo de terapia anti-retroviral (NR: não reportado).

3.2. Amplificação por RT-nested PCR

Como referido no capítulo anterior, neste estudo, a amplificação da região C2V3C3 de gp120 do gene *env* de HIV-1 foi realizada por transcrição reversa *in vitro* seguida de *nested* PCR (RT-nested PCR), a partir de 52 amostras de RNA viral purificado a partir do plasma. Foi possível obter um produto de amplificação, com o tamanho esperado de 565 pb, para 35 destas amostras (67%). Um resultado característico está representado na **Figura 13**.

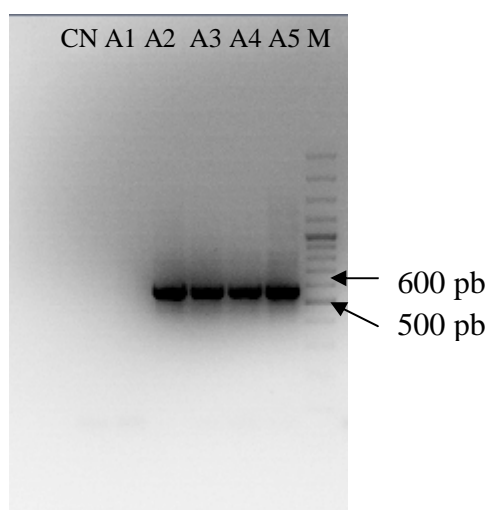


Figura 13: Imagem digital captada de um gel de agarose a 1,5% (p/v), em tampão TAE 0,5X, com brometo de etídeo (0,5 µg/ml), após separação electroforética, onde se pode observar o resultado da amplificação por RT-*nested* PCR do RNA viral de várias amostras. CN - Controlo negativo (ou de contaminação); A1 – Amostra PT_TX20; A2 – Amostra PT_TX37; A3 – Amostra PT_TX50; A4 – Amostra PT_TX59; A5 – Amostra PT_TX60; M – Marcador de massa molecular *Gene Ruler™ 100bp Plus DNA Ladder* (Fermentas, Canadá).

As condições de amplificação foram seleccionadas de acordo com os sistemas comerciais utilizados, considerando as temperaturas de hibridação estimadas para os *primers* utilizados. No entanto, a obtenção de fragmentos específicos nem sempre foi conseguida, pelo que foi necessário recorrer à optimização das condições de amplificação, processo de especial importância para se garantir a amplificação do fragmento correcto, sem produtos inespecíficos. Durante este processo, testaram-se diferentes temperaturas de hibridação dos *primers*, quer para a fase de retrotranscrição (55, 50 e 45°C), quer para a 1ª PCR (55, 50, 45 e 40°C). Foi também testado o número de ciclos da fase de amplificação da 1ª PCR e a utilização, ou não, de incrementos de tempo na fase de polimerização. Por fim, testou-se igualmente a utilização de um par de *primers* para a retrotranscrição e 1ª PCR ou a utilização, em simultâneo, de dois pares de *primers* (**Tabela 2**). A segunda reacção de amplificação (2ª PCR) não foi sujeita a qualquer optimização.

De um total de 35 produtos amplificados, foi possível proceder-se à purificação directa a partir dos produtos de PCR e envio imediato para sequenciação nucleotídica de 21 (60%). Relativamente aos restantes, passou-se por um processo intermédio de clonagem no vector plasmídico *pGEM®-T Easy* (Promega, EUA), de acordo com o protocolo descrito no capítulo anterior.

Por fim, pesquisou-se, por métodos estatísticos, a eventual existência de associação entre o sucesso/insucesso de amplificação e os dados epidemiológicos dos indivíduos dadores das amostras estudadas (**Anexo 2**). Através desta análise, verificou-se não existir qualquer influência das características epidemiológicas ou clínicas conhecidas no sucesso da amplificação, uma vez que se pode observar uma distribuição muito semelhante das diferentes frequências entre as amostras amplificadas e não amplificadas, com excepção da categoria “terapia anti-retroviral”. Neste caso, o teste exacto de Fisher bilateral, aplicado à análise de uma tabela de contingência 2x2, permitiu a rejeição da hipótese de independência das duas variáveis ($p < 0,001$).

3.3. Edição das Sequências Nucleotídicas e Pesquisa de Homologia Genética

Após a amplificação, os fragmentos de cDNA foram enviados para sequenciação nucleotídica em ambas as cadeias, desta forma tentando minimizar-se a introdução de erros devidos à própria sequenciação. As sequências nucleotídicas de cada uma das cadeias, dos 35 fragmentos amplificados, foram editadas, por comparação das cadeias simples, directa e inversa complementar, duas a duas. Através de uma cuidadosa análise dos cromatogramas respectivos, realizou-se a correcção manual das discrepâncias e, nas posições onde não foi possível chegar a um nucleótido “consenso”, utilizou-se o código degenerado de nucleótidos. Para várias sequências, verificou-se uma elevada variabilidade, relativamente a determinadas posições, reflectindo a existência de *quasispecies*.

Uma vez editadas, as sequências nucleotídicas foram utilizadas para uma pesquisa automática de homologia genética, utilizando o programa *BLAST* do NCBI (*National Center for Biotechnology Information*, EUA), a fim de confirmar a sua identidade, prosseguindo-se depois para a análise filogenética.

3.4. Análise Filogenética

Primeiramente, procedeu-se a uma genotipagem automática, com recurso ao programa *REGA HIV-1 & 2 Automated Subtyping Tool v.2.0* (disponível em <http://www.bioafrica.net/rega-genotype/html/subtypinghiv.html>). Das 35 amostras estudadas, 65,7% (23) foram classificadas do subtipo B de HIV-1, 22,9% (8) do subtipo A, 8,6% (3) do subtipo G e 2,9% (1) do sub-subtipo F1. Estes resultados foram

posteriormente confirmados através da construção não-automatizada de árvores filogenéticas, o que permitiu a selecção, entre outros, dos algoritmos e das sequências de referência a utilizar, constituindo, assim, uma metodologia mais abrangente.

A árvore filogenética foi construída com base no alinhamento múltiplo das 35 sequências em estudo com 31 sequências de referência, com recurso ao programa *ClustalW2* v.2.0. (disponível em <http://www.ebi.ac.uk/Tools/msa/clustalw2/>). O alinhamento englobava as sequências amplificadas, sequências de referência representativas do grupo M (subtipos A-D, F-H, J e K, sub-subtipos A1-A4 e F1-F2 e formas recombinantes circulantes CRF02_AG e CRF14_BG) e uma sequência de SIV, como *outgroup* (com origem na base de dados de Los Alamos, em <http://www.hiv.lanl.gov/>). Após edição cuidadosa, com recurso ao programa *BioEdit* v.7.0.9.0., o alinhamento foi utilizado para a construção da árvore filogenética através do programa *MEGA* v.4.0 (disponível em <http://www.megasoftware.net/>). Utilizou-se o método de *neighbor-joining* para a construção da árvore, implementado através do modelo de 2 parâmetros de Kimura (91) para o cálculo da matriz de distâncias genéticas. Foi também utilizado o método de *bootstrapping*, com base em 1000 amostragens aleatórias, para avaliar a robustez da inferência filogenética e das topologias sugeridas. Obteve-se então uma árvore filogenética com 66 sequências, com base num alinhamento múltiplo de 454 nucleótidos (**Figura 14**).

Numa análise preliminar da árvore filogenética pode verificar-se que todas as sequências estudadas são suportadas por ramos horizontais, de tamanho variável, o que demonstra divergência em relação a um ancestral comum, indicando que não existem amostras com sequências nucleotídicas iguais entre si, permitindo excluir a ocorrência de contaminações durante o protocolo experimental.

A análise filogenética realizada para as sequências nucleotídicas correspondentes à região amplificada permitiu a identificação de 23 sequências do subtipo B, sendo 2 destas do tipo da CRF14_BG, distribuídas em 6 subgrupos distintos, dentro do agrupamento que engloba todas as sequências do subtipo B (*bootstrap* de 81%). O primeiro subgrupo, que não engloba nenhuma sequência de referência, compreende 5 sequências (PT_TP184A, PT_TX16, PT_TP107, PT_TP71 e PT_TX40) e apresenta um ancestral comum com outro subgrupo que compreende as sequências *env* B presentes na CRF14_BG. Neste, encontram-se 2 sequências anteriormente

subtipadas como B (PT_TX32 e PT_TX45) e 2 sequências de referência da CRF14_BG. Pode inferir-se que as sequências PT_TX32 e PT_TX45 pertencem à forma recombinante circulante CRF14_BG ou que são evolutivamente aparentadas com as estirpes B suas parentais. Num terceiro subgrupo, encontram-se as sequências PT_TP02 e PT_TP159A agrupadas com a sequência de referência B_DO_05DO 160884. As sequências PT_TX18, PT_TP16 e PT_TX60 encontram-se agrupadas, formando um novo subgrupo, com as sequências de referência B_FR_HXB2 e B_YE_02YE507. Num quinto subgrupo, encontram-se agrupadas, com a sequência de referência B_BO_BOL0122, 3 sequências (PT_TP114, PT_TX47 e PT_TP85). As restantes 8 sequências em estudo do subtipo B (PT_TX30, PT_TP24, PT_TP38, PT_TX21, PT_TP219, PT_TX50, PT_TX37 e PT_TX59) encontram-se agrupadas com a sequência de referência B_DK_PMVL 012. Digno de nota que, exceptuando o agrupamento que contém as sequências do tipo CRF14_BG, nenhum dos agrupamentos mencionados é suportado por valores significativos de *bootstrap*.

A análise da árvore filogenética (**Figura 14**) revela ainda uma estirpe pertencente ao subtipo F (PT_AS646), mais especificamente ao sub-subtipo F1, e 3 pertencentes ao subtipo G (PT_TP98, PT_TP13 e PT_TX14) que se encontram agrupadas com duas sequências de referência do subtipo G, uma de Portugal e outra de Espanha. Pode-se ainda identificar 8 sequências (PT_TX46, PT_TX13, PT_TX41, PT_TX15, PT_TP19, PT_TP208, PT_TP156 e PT_TX12) classificadas como subtipo A, agrupadas, embora de um modo não significativo, com diversas sequências de referência dos diferentes sub-subtipos de A (A1-A4) e CRF02_AG, forma recombinante circulante com *env* do subtipo A.

Os resultados de genotipagem obtidos através da análise filogenética manual estão de acordo com os da pesquisa rápida implementada pelo programa *REGA HIV-1 & 2 Automated Subtyping Tool v.2.0*. No caso das sequências que na árvore aparecem agrupadas com as formas recombinantes circulantes CRF14_BG (PT_TX32 e PT_TX45), estas foram anteriormente subtipadas como B, dado que a estrutura mosaico de CRF14_BG corresponde, na região de *env* estudada, a esse mesmo subtipo (40). Através da análise filogenética pode então inferir-se que, na globalidade, 21 sequências são do subtipo B (60%) e 14 de subtipos não-B (40%), mais especificamente, A, G, F e CRF14_BG. Na **Figura 14**, o agrupamento B apresenta uma elevada

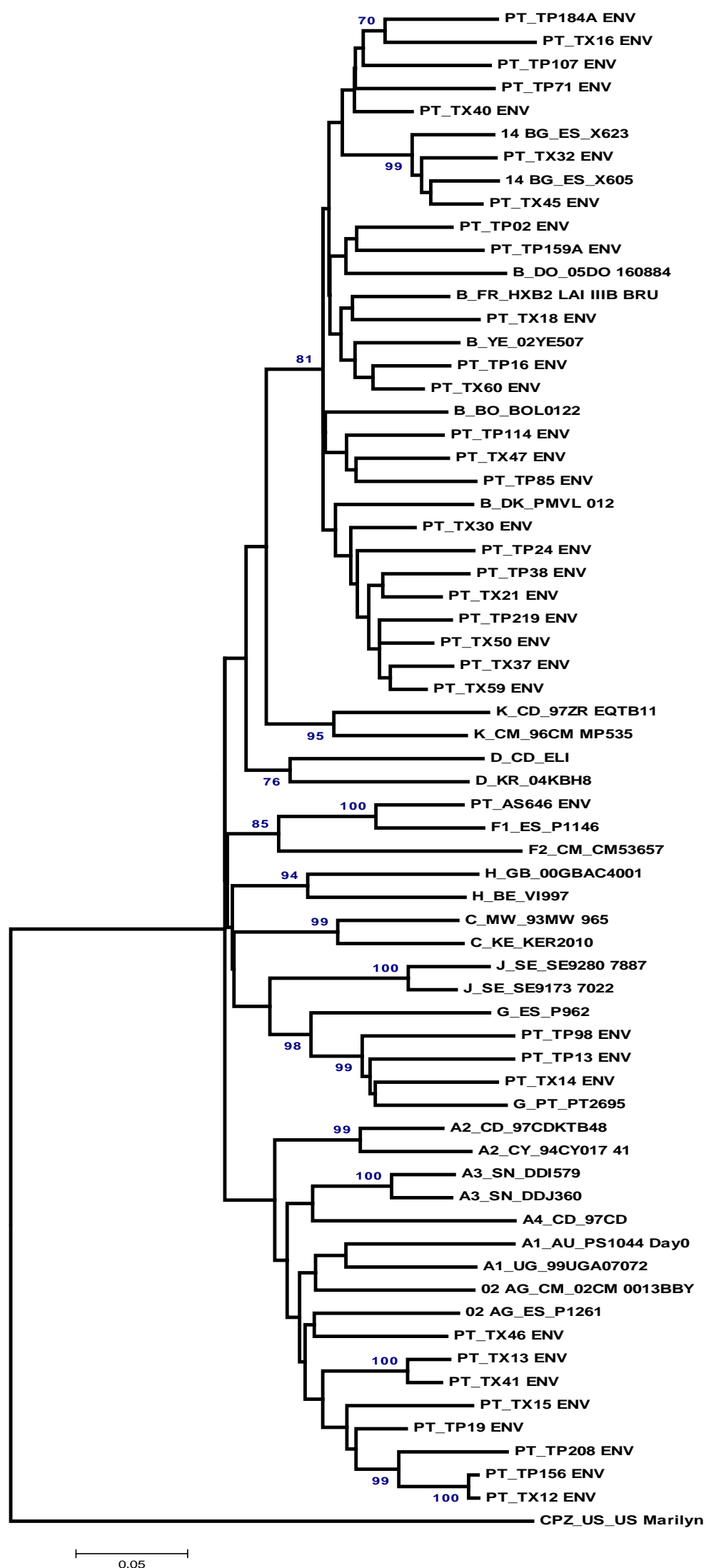


Figura 14: Resultado da análise filogenética realizada à região C2V3C3 do gene *env* de HIV-1 das 35 amostras em estudo (representadas com o prefixo PT), com 30 sequências de referência do grupo M (subtipos A-D, F-H, J e K, sub-subtipos A1-A4 e F1-F2 e as CRF02_AG e 14_BG) e a sequência de SIV CPZ_US_US Marilyn, utilizada como *outgroup* (sequências retiradas de <http://hiv-web.lanl.gov/>), a partir de um alinhamento múltiplo de 454 nucleótidos. Na designação de cada sequência de referência, o prefixo corresponde ao subtipo (ou CRF) seguido da abreviatura do país do isolado e por fim o nome do isolado viral. A árvore foi construída com o programa *MEGA v.4.0*, utilizando-se o método de *neighbor-joining*, baseado no modelo de 2 parâmetros de Kimura para o cálculo da matriz de distâncias genéticas (a escala indica 5% de divergência). A robustez da inferência filogenética foi testada por *bootstrapping* de 1000 replicados, sendo representados apenas os valores considerados significativos ($\geq 70\%$).

diversidade genética, ocorrendo a formação de 6 subgrupos distintos, mas apenas o subgrupo que contém as sequências CRF14_BG apresenta um valor de *bootstrap* significativo (99%). Por seu lado, o agrupamento das sequências do subtipo G é suportado por um valor significativo de *bootstrap* (98%), assim como o agrupamento das sequências F (85%). O mesmo não acontece, no entanto, para o subtipo A (valor de *bootstrap* inferior a 70%, que não é um valor significativo). De realçar que quase todas as sequências portuguesas, à excepção de PT_TX46, se agrupam à parte, relativamente às referências seleccionadas para a construção da árvore filogenética, embora com um valor de *bootstrap* não significativo. Por fim, a sequência PT_AS646 agrupa com a sequência de referência F1, com um valor de *bootstrap* de 100%, pelo que foi considerada como pertencente a este sub-subtipo.

A análise filogenética permitiu distribuir as sequências da amostra entre os subtipos A, B, G, sub-subtipo F1 e CRF14_BG. Esta distribuição está representada na **Figura**

15, onde se pode verificar que o subtipo prevalente é o B, que

engloba 60% das sequências, e o menos frequente é o sub-subtipo F1, com apenas 2,9%. Foi também realizada uma análise de *bootscanning* para as 35 sequências, não se tendo observado quaisquer indícios de recombinação intergenotípica (resultados não apresentados).

Uma segunda árvore filogenética foi construída (**Figura 16**), integrando as 35 sequências em estudo, 22 sequências de referência, extraídas da base de dados de Los Alamos, e um conjunto de 33 sequências do gene *env* de HIV-1 de origem portuguesa, obtidas na base de sequências nucleotídicas do NCBI. Com base num alinhamento múltiplo de 419 nucleótidos, a construção desta árvore teve como objectivo verificar o tipo de relação filogenética existente entre as sequências estudadas no decurso deste

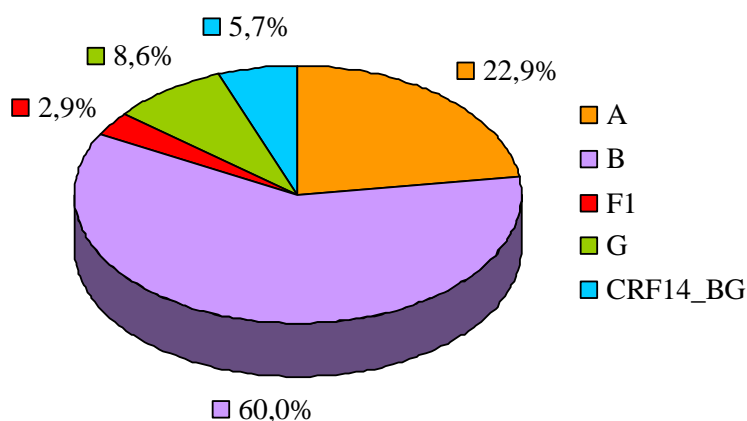


Figura 15: Representação gráfica da distribuição dos subtipos de HIV-1 presentes na amostra.

trabalho e outras sequências reportadas anteriormente de Portugal. Através da sua análise, foi possível observar a ausência de qualquer tipo de segregação, nomeadamente, de tipo temporal, entre as sequências portuguesas previamente estudadas (de proveniências diversas) e as novas, havendo uma mistura entre ambos os tipos e entre aquelas e as referências utilizadas. Observa-se a existência de divergências genéticas relativamente elevadas (**Figura 16**), o que era esperado uma vez que está a ser estudada uma das regiões mais variáveis do genoma de HIV-1. Na generalidade, mantêm-se as topologias da árvore filogenética da **Figura 14**, continuando a ser evidentes alguns subgrupos, embora não significativos, no grupo de sequências do subtipo B (na globalidade, estas sequências foram agrupadas com um valor de *bootstrap* significativo de 84%). Destes agrupamentos, o único com significância (*bootstrap* de 99%) contém sequências de referência da forma recombinante circulante CRF14_BG, 2 sequências em estudo (PT_TX32 e PT_TX45) e 3 outras sequências portuguesas (BM23439, BM23061 e BM21342). Todos os restantes agrupamentos mantêm-se genericamente consistentes com os resultados obtidos anteriormente, reeditando-se os emparelhamentos já conhecidos entre sequências em estudo, com intercalação das sequências portuguesas provenientes da base de dados. A única exceção prende-se com a divisão do agrupamento contendo as sequências PT_TP02, PT_TP159A e sequência de referência B_DO_05DO (**Figura 14**) em dois subgrupos bem distintos, um dos quais contendo a sequência de origem portuguesa BM23440 (**Figura 16**).

As sequências do subtipo G encontram-se agrupadas em conjunto, com um valor elevado de *bootstrap* (96%). Este grupo subdivide-se em dois subgrupos consistentes (97% e 90% de *bootstrap*) e, curiosamente, o mais significativo engloba apenas sequências de origem portuguesa. O agrupamento das sequências do subtipo F apresenta, como era antecipado, um subgrupo composto apenas de sequências do subsubtipo F1, com um valor de *bootstrap* de 99%, contendo uma sequência de referência de origem espanhola destacada e um grupo de sequências que integra a PT_AS646 e três sequências portuguesas já estudadas (BM24871, PT955i e 103). Finalmente, as sequências do subtipo A encontram-se num grande grupo, estatisticamente pouco significativo, à imagem do resultado da árvore filogenética da **Figura 14**, incluindo sequências de referência dos sub-subtipos A1 a A4, CRF02_AG (*env* A) e as sequências de origem portuguesa, observando-se a ausência de qualquer estruturação. Tal como

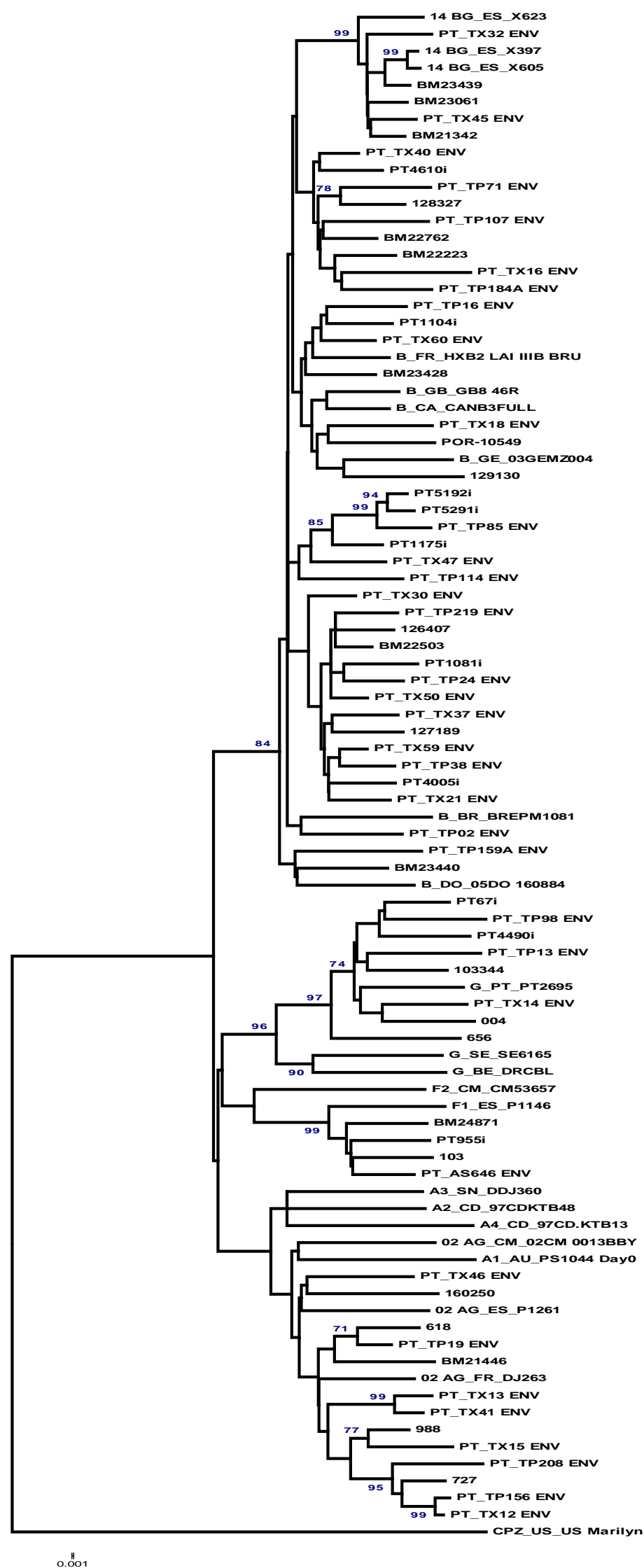


Figura 16: Resultado da análise filogenética realizada à região C2V3C3 do gene *env* de HIV-1 das 35 amostras em estudo (representadas com o prefixo PT), com 33 sequências de origem portuguesa, estudadas anteriormente e depositadas na base de dados do NCBI (várias designações), 21 sequências de referência do grupo M (subtipos A-D, F-H, J e K, sub-subtipos A1-A4 e F1-F2 e as CRF02_AG e 14_BG) e a sequência de SIV CPZ_US_US_Marylin, utilizada como *outgroup* (<http://hiv-web.lanl.gov/>), a partir de um alinhamento múltiplo de 419 nucleótidos. Na designação de cada sequência de referência, o prefixo corresponde ao subtipo (ou CRF), seguido da abreviatura do país do isolado e por fim o nome do isolado viral. A árvore foi construída com o programa *MEGA v.4.0*, utilizando-se o método de *neighbor-joining*, baseado no modelo de 2 parâmetros de Kimura para o cálculo da matriz de distâncias genéticas (a escala indica 0,1% de divergência). A robustez da inferência filogenética foi testada por *bootstrapping* de 1000 replicados, sendo representados apenas os valores considerados significativos ($\geq 70\%$).

anteriormente, a sequência PT_TX46, agora emparelhada com a sequência portuguesa 160250, aparenta maior proximidade filogenética, embora pouco significativa, com a forma genética A do gene *env* da CRF02_AG.

3.5. Análise das Sequências de Aminoácidos

As sequências inferidas de aminoácidos foram analisadas recorrendo ao seu alinhamento múltiplo, após a tradução das sequências nucleotídicas respectivas, com a sequência consenso do subtipo B, retirada da base de dados de Los Alamos (**Figura 17**). As sequências em estudo (n=35) estão organizadas de acordo com o seu subtipo, inferido da análise filogenética anteriormente realizada, encontrando-se dispostas por ordem alfabética: PT_TP19 a PT_TX46, subtipo A (n=8); PT_TP02 a PT_TX60, subtipo B (n=23); PT_AS646, sub-subtipo F1 (n=1); PT_TP13 a PT_TX14, subtipo G (n=3). Na **Figura 17**, foi também assinalada a região da ansa V3 (aa 91-125) da gp120.

As sequências apresentam um número variável de aminoácidos, compreendido entre 168 e 172, com cerca de metade (n=17) a conter 169 aminoácidos, à imagem da sequência consenso B utilizada no alinhamento. Ocasionalmente, foram abertos espaços de modo a manter o alinhamento, incluindo em duas posições da sequência de referência (Cons_B), o que poderá ser interpretado como ocorrência de inserções e deleções pontuais. A merecer destaque, a sequência PT_TP02 com 172 aminoácidos devido a uma inserção de 2 dois resíduos, um de treonina e outro de serina, entre os aminoácidos 82 e 83 da sequência consenso B (posições 83-84 do alinhamento da **Figura 17**). Não se observaram inserções ou deleções de grande extensão, nem a ocorrência de codões precoces de terminação da tradução. A extensa variabilidade genética, característica da região do genoma do HIV-1 estudada, teve como consequência a impossibilidade de determinar, sem ambiguidade, a natureza dos aminoácidos em 96 posições distribuídas por 14 das sequências estudadas (posições assinaladas com X, na **Figura 17**). No entanto, dada a relevância desta informação para a análise de polimorfismos genéticos e sua associação à resistência ao maraviroc (ver mais à frente), foram estudadas as diferentes hipóteses possíveis de tradução para cada uma das posições nucleotídicas ambíguas, sendo os resultados obtidos apresentados na **Tabela 5**.

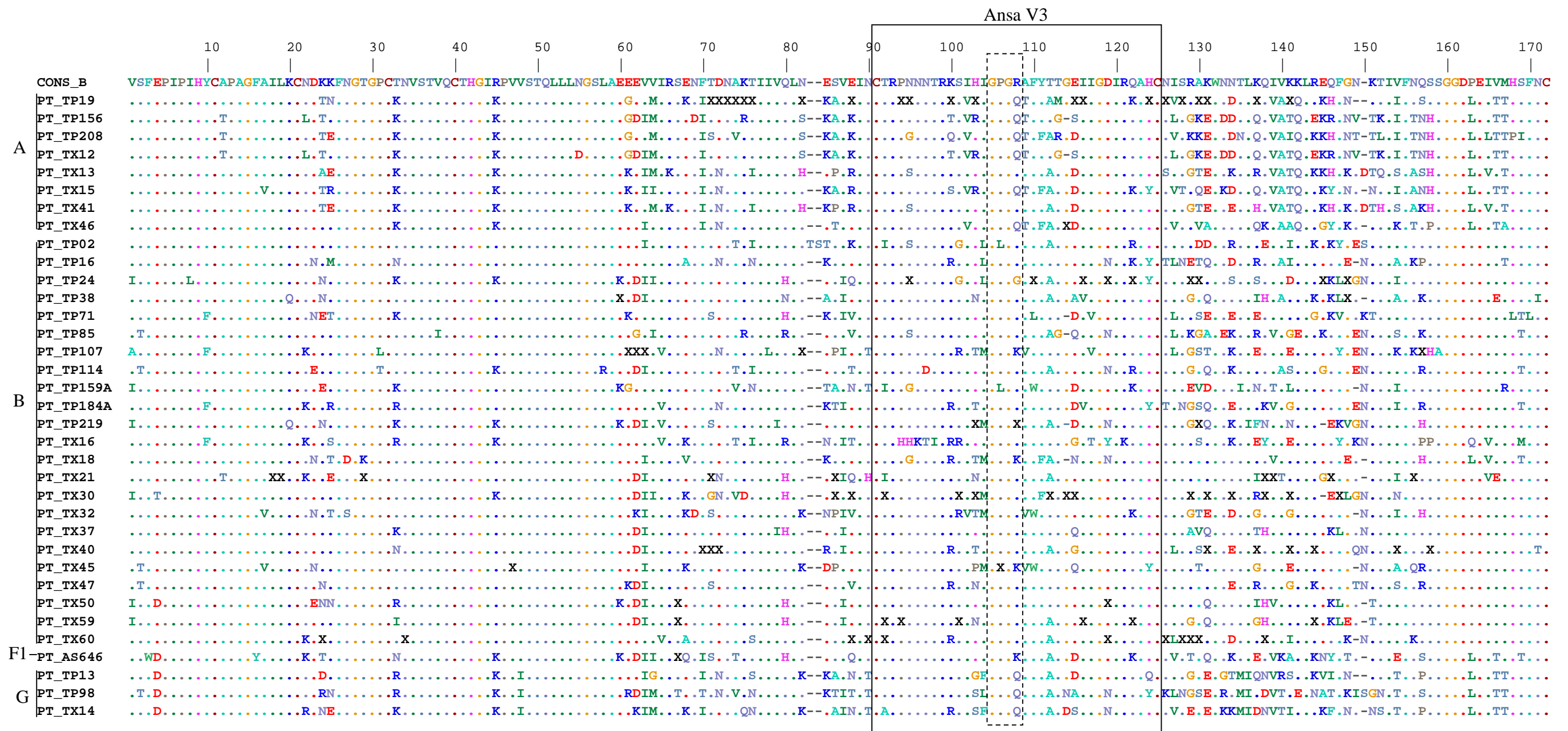


Figura 17: Alinhamento das sequências de aminoácidos da região C2V3C3 do gene *env* de HIV-1 para as amostras estudadas. As sequências são designadas pelo prefixo PT (Portugal) seguido pela referência da amostra, encontrando-se alinhadas com a sequência consenso do subtipo B (CONS_B) retirada da base de dados de Los Alamos (<http://www.hiv-web.lanl.gov/>). As sequências são compostas por letras de cores diferentes correspondentes a diferentes tipos de aminoácidos, utilizando o código universal de uma letra. Os pontos significam identidade dos aminoácidos com a sequência consenso, enquanto os traços indicam inserções ou deleções. A letra “X” representa a possibilidade de mais de um aminoácido numa determinada posição (**Tabela 5**), devido à utilização anterior do código degenerado de nucleótidos. A região da ansa V3 encontra-se delimitada por um rectângulo (posições 91 a 125 do alinhamento). A coroa da ansa V3 encontra-se delimitada por um rectângulo menor, a tracejado (posições 105 a 108 do alinhamento).

Tabela 5: Apresentação das diferentes hipóteses de tradução para cada posição nucleotídica na qual foi utilizado o código degenerado de nucleótidos, devido à presença de *quasispecies* (todas as posições representadas com X no alinhamento de aminoácidos da **Figura 17**).

Sequência	Polimorfismos	Sequência	Polimorfismos	Sequência	Polimorfismos
PT_TP19	71T/S	PT_TP24	93S/N	PT_TX40	70L/F
	72D/N		108L/F		71T/S
	73S/I		114V/L		72N/D
	74G/V		117D/N		129K/R
	75Y/*S/P/H/Q		120W/*R		135K/E
	76N/T		127G/V/R/I		139E/G
	82S/N		128A/T		142R/K
	86R/K		143E/K		152V/Y
	92Y/H		146G/E		156S/P
	93S/N	PT_TP219	101N/T	PT TX59	67A/S
	98E/K		106R/K		90I/T
	101R/H		127A/T		92T/P
	113E/D	PT_TX30	84T/S		99G/S
	114V/I		86K/E		114V/I
	122Q/H		90I/V		120R/Q
	124D/H/N		99S/G		142R/K
	126S/N		101H/P/Y/S	PT TX60	24E/K
	128R/K		110T/A		34E/K
	129E/D		112Q/R/E/G		86E/V/K/I
	135W/*R		113G/E		88E/D/K/N
PT_TX21	139A/T		127K/R/E/G		90V/A/I/T
	18I/L		129K/E		117E/D/K/N
	19I/L		132K/N		124N/T/I
	29K/T		136H/Q		126S/N
	71K/N/T		139K/N/R/S		127G/R
	84A/S		144K/Q		128A/T
	136Q/R/H	PT_TX45	47V/A		136Q/H
	137I/V		104T/P	PT_TP107	61E/K
	144K/E	PT TX50	67F/S		62G/E
	153T/S		117D/N		63V/I
PT_AS646	67S/F	PT_TP38	60K/Q		82N/H
PT_TX46	112G/D		146G/A/E		155E/K

[A numeração corresponde à posição na respectiva sequência. As letras representam os diferentes aminoácidos, utilizando o código universal de uma letra. O símbolo * representa a introdução de um codão *stop* aquando da tradução.]

Em termos genéricos, a região mais heterogênea do alinhamento, no que se refere à natureza dos aminoácidos, corresponde à localizada imediatamente a jusante da ansa V3 (a partir da posição 126).

Relativamente à ansa V3, verifica-se que 6 sequências (PT_TP71, PT_TP85, PT_TP156, PT_TP219, PT_TX12 e PT_TX18) apresentam uma deleção na posição 114 do alinhamento, sendo a sua ansa V3 composta por 34 aminoácidos, enquanto todas as restantes (n=29; 82,9%) têm 35 aminoácidos, tal como a sequência consenso do subtipo B. Pode-se também verificar que todas as sequências exibem as posições 91C e 125C conservadas (havendo mais 7 posições nestas circunstâncias: 93R, 99R, 105G, 107G, 118G, 120I e 123A), correspondendo aos resíduos de cisteína que mantêm a conformação tridimensional da ansa V3 através de uma ponte dissulfeto (daí a importância da sua conservação). Observaram-se 5 diferentes motivos para a coroa da ansa V3, com o motivo GPGR, presente no consenso B, em maior número de sequências (n=18, sendo 2 do subtipo A e 16 do subtipo B). O motivo GPGQ, segundo mais frequente (n=9, 6 A e 3 G), é claramente o mais comum em sequências não-B. Foram também encontrados os motivos GLGR, presente em duas sequências do subtipo B, o motivo GPGG, presente em apenas uma sequência B, e o motivo GPGK, presente em 3 sequências (2 B e F1). Para duas sequências (ambas do subtipo B) verificou-se a existência de ambiguidades de tradução numa das posições do motivo (PT_TP219: GPGR/K; PT_TX45: GP/TGK), pelo que estas foram excluídas da análise estatística realizada. Recorrendo, uma vez mais, ao teste exacto de Fisher bilateral, aplicado à análise de uma tabela de contingência 5x2, foi possível rejeitar, com $p < 0,001$, a hipótese de independência entre o genótipo das sequências (B vs. não-B) e os motivos aminoacídicos presentes na coroa da ansa V3, havendo uma associação entre o motivo GPGR e o subtipo B e entre o motivo GPGQ e os subtipos não-B (neste estudo, A e G).

3.5.1. Análise de Polimorfismos Genéticos

A comparação das sequências de aminoácidos obtidas com a sequência consenso B demonstrou a existência de um elevado número de polimorfismos naturais na região estudada, tal como era esperado, dado tratar-se de uma das regiões mais variáveis do genoma do HIV-1. Através da análise da **Figura 18**, pode observar-se que das 172

										5					10					15					20				
V ₈₀	S _{85,7}	F _{97,1}	E _{82,9}	P ₁₀₀	I ₁₀₀	P ₁₀₀	I _{97,1}	H ₁₀₀	Y _{88,6}	C ₁₀₀	A _{88,6}	P ₁₀₀	A ₁₀₀	G ₁₀₀	F _{97,1}	A _{91,4}	I _{98,6}	L _{98,6}	K _{94,3}	C ₁₀₀	N _{74,3}								
I _{17,1}	T _{14,3}	W _{2,9}	D _{14,3}				L _{2,9}		F _{11,4}		T _{11,4}				Y _{2,9}	V _{8,6}	L _{1,4}	I _{1,4}	Q _{5,7}		K _{17,1}								
A _{2,9}			T _{2,9}																		L _{5,7}								
																					R _{2,9}								
										30					35					40									
D ₈₀	K _{47,1}	K _{57,1}	F ₁₀₀	N _{94,3}	G ₁₀₀	T _{95,7}	G ₁₀₀	P _{94,3}	C ₁₀₀	T _{34,3}	N _{97,1}	V ₁₀₀	S ₁₀₀	T ₁₀₀	V _{97,1}	Q ₁₀₀	C ₁₀₀	T ₁₀₀	H ₁₀₀	G ₁₀₀	I ₁₀₀								
N _{14,3}	T ₂₀	E _{14,3}		D _{2,9}		K _{4,3}		L _{2,9}		K ₄₀	E _{1,4}				I _{2,9}														
E _{5,7}	N _{17,1}	N _{8,6}		S _{2,9}				T _{2,9}		R _{14,3}	K _{1,4}																		
	E _{7,1}	T _{8,6}								N _{8,6}																			
	A _{2,9}	R _{5,7}								I _{2,9}																			
	D _{2,9}	M _{2,9}																											
	R _{2,9}	S _{2,9}																											
										50					55					60					65				
R _{51,4}	P ₁₀₀	V _{98,6}	V _{91,4}	S ₁₀₀	T ₁₀₀	Q ₁₀₀	L ₁₀₀	L ₁₀₀	L ₁₀₀	N _{97,1}	G ₁₀₀	S ₁₀₀	L _{97,1}	A ₁₀₀	E _{82,9}	E _{67,1}	E _{47,1}	V ₃₀	V ₆₀	I _{85,7}	R _{94,3}								
K _{48,6}		A _{1,4}	I _{8,6}							D _{2,9}			R _{2,9}		K _{15,7}	K _{15,7}	D _{42,9}	I ₇₀	M _{22,9}	V _{14,3}	K _{5,7}								
															Q _{1,4}	G _{14,3}	K _{5,7}		I _{14,3}										
																R _{2,9}	G _{4,3}		G _{2,9}										

^ 70										75										80										^ 85									
S _{92,9}	E _{71,4}	N _{94,3}	F _{64,3}	T _{75,1}	D _{65,7}	N _{97,1}	A _{74,3}	K _{85,7}	T _{67,1}	I ₁₀₀	I _{97,1}	V _{94,3}	Q _{68,6}	L ₁₀₀	N _{71,4}	Δ _{97,1}	Δ _{97,1}	E _{45,7}	S _{51,4}	V ₆₀	E _{36,4}																		
F _{2,9}	K _{17,1}	D _{5,7}	I _{31,4}	S ₂₀	N _{34,3}	I _{1,4}	V _{12,9}	R _{5,7}	I _{14,3}		L _{2,9}	I _{5,7}	H _{22,9}		K _{11,4}	T _{2,9}	S _{2,9}	K _{34,3}	A _{24,3}	I ₄₀	K ₁₅																		
T _{2,9}	A _{5,7}		T _{2,9}	G _{2,9}		S _{1,4}	T _{11,4}	Q _{3,3}	N _{12,9}				R _{5,7}		S ₁₀			N _{5,7}	P _{14,3}		V _{12,1}																		
A _{1,4}	Q _{2,9}		L _{1,4}	K _{0,9}			G _{1,4}	D _{2,9}	S _{5,7}				N _{2,9}		H _{7,1}			T _{5,7}	T ₁₀		R ₁₀																		
	V _{2,9}			N _{0,9}				H _{0,5}										A _{2,9}			N _{8,6}																		
								P _{0,5}										D _{2,9}			Q _{8,6}																		
								S _{0,5}										R _{2,9}			T _{8,6}																		
								Y _{0,5}													I _{0,7}																		
								* _{0,5}																															

90^										95 ^										100										105										110									
I ₁₀₀	N _{80,7}	C ₁₀₀	T _{82,1}	R ₁₀₀	P _{92,9}	N _{74,3}	N _{97,1}	N _{94,3}	T _{97,1}	R ₁₀₀	K _{64,3}	S _{82,9}	I ₈₀	H _{47,9}	I _{65,7}	G ₁₀₀	P _{92,9}	G ₁₀₀	R _{58,6}	A _{74,3}	F _{87,1}																												
	T _{14,3}		I _{12,1}		H _{4,3}	S _{14,3}	K _{2,9}	D _{2,9}	I _{2,9}		R _{22,9}	G _{8,6}	V ₂₀	T _{15,7}	M _{17,1}		L _{5,7}		Q _{25,7}	T _{17,1}	W _{8,6}																												
	H _{2,9}		A _{3,6}		T _{1,4}	G _{8,6}		T _{2,9}			T _{5,7}	R _{8,6}		N _{12,9}	L _{11,4}		T _{1,4}		K _{12,9}	V _{8,6}	L _{4,3}																												
	D _{0,7}		V _{2,1}		Y _{1,4}	H _{2,9}					Q _{2,9}			R ₁₀	F _{5,7}				G _{2,9}																														
	E _{0,7}										S _{2,9}			S _{6,4}																																			
	K _{0,7}										E _{1,4}			P _{3,6}																																			
															G _{2,9}																																		
														Y _{0,7}																																			

										115					120					125					130									
Y _{85,7}	T _{38,6}	T _{85,7}	G _{73,6}	E _{34,3}	I _{88,6}	I _{91,4}	G ₁₀₀	D _{72,1}	I ₁₀₀	R _{97,1}	Q _{72,9}	A ₁₀₀	H _{78,6}	C ₁₀₀	N _{84,6}	I _{57,1}	S _{85,7}	R _{37,1}	A _{35,7}	K _{22,9}	W ₁₀₀													
F _{14,3}	A _{61,4}	G _{8,6}	Δ _{17,1}	D _{32,9}	V ₁₀	V _{5,7}		N _{23,6}		K _{2,9}	K _{17,1}		Y _{17,1}		T _{6,6}	L _{25,7}	N _{11,4}	G ₄₀	T _{18,6}	E _{34,3}														
		M _{2,9}	D _{4,3}	Q _{8,6}	L _{1,4}	T _{2,9}		Y _{2,9}			R ₈		Q _{4,3}		K _{2,9}	V _{17,1}	T _{2,9}	E _{9,3}	S _{17,1}	Q _{25,7}														
		R _{2,9}	N _{2,9}	S _{8,6}				E _{0,7}			W _{0,9}				S _{2,9}			K _{6,4}	K ₁₀	D _{7,1}														
			E _{0,7}	G _{7,1}				K _{0,7}			* _{0,9}				D _{0,9}			A _{2,9}	V _{8,6}	A _{5,7}														
			Q _{0,7}	A _{5,7}											H _{0,9}			T _{2,9}	D _{2,9}	T _{2,9}														
			R _{0,7}	N _{2,9}											I _{0,9}			I _{0,7}	G _{2,9}	R _{1,4}														
																		V _{0,7}	Q _{2,9}															
																			R _{1,4}															
					135					140					145										150									
N ₈₀	N ₃₀	T _{85,7}	L _{88,6}	K _{18,6}	Q _{60,9}	I _{52,9}	V _{54,3}	K _{23,6}	K _{71,4}	L ₁₀₀	R _{65,7}	E ₅₀	Q _{37,1}	F _{68,6}	G _{45,1}	N _{37,1}	Δ ₄₀	K _{68,6}	T ₈₀	I ₁₀₀	V ₃₀													
D _{5,7}	K _{24,3}	M _{8,6}	I _{11,4}	Q _{17,1}	H _{15,1}	V _{44,3}	A _{31,4}	T _{15,7}	Q _{22,9}		K _{22,9}	K _{27,1}	K _{21,4}	L _{14,3}	N ₂₀	Δ ₂₀	N _{45,7}	T _{22,9}	K _{5,7}		I _{28,6}													
K _{5,7}	D _{22,9}	I _{2,9}		R _{15,1}	N _{8,6}	T _{2,9}	T _{8,6}	I _{14,3}	E _{5,7}		E _{5,7}	G _{8,6}	H _{11,4}	Y _{8,6}	K _{14,3}	E _{14,3}	D _{5,7}	N _{5,7}	H _{2,9}		S _{11,4}													
E _{2,9}	E _{14,3}	N _{2,9}		E _{12,9}	K _{5,7}		K _{2,9}	G _{12,9}			G _{2,9}	Δ _{5,7}	Y _{8,6}	K _{4,3}	E _{10,9}	G _{8,6}	S _{5,7}	G _{2,9}	L _{2,9}		A _{8,6}													
G _{2,9}	R _{2,9}			G _{8,6}	D _{2,9}		R _{2,9}	E ₁₀			N _{2,9}	A _{2,9}	E _{7,1}	V _{2,9}	L _{2,9}	T _{5,7}	K _{2,9}		N _{2,9}		T _{8,6}													
R _{2,9}	S _{2,9}			I _{8,6}	E _{2,9}			A _{7,1}				N _{2,9}	R _{5,7}	Q _{1,4}	T _{2,9}	V _{5,7}			Q _{2,9}		K _{5,7}													
	T _{2,9}			D _{2,9}	Y _{2,9}			S _{6,4}				V _{2,9}	F _{2,9}		V _{2,9}	I _{2,9}			S _{2,9}		E _{2,9}													
				F _{2,9}	R _{0,9}			N _{3,6}					I _{2,9}		A _{0,9}	K _{2,9}					N _{2,9}													
				H _{2,9}				D _{2,9}					T _{2,9}			Q _{2,9}					Y _{1,4}													
				N _{2,9}				L _{2,9}																										
				S _{2,9}				R _{0,7}																										
				T _{2,9}																														
				W _{0,9}																														
				* _{0,9}																														

155	^				160						165					170			
F ₁₀₀	N _{65,7}	Q _{34,3}	S _{72,9}	S _{97,1}	G ₁₀₀	G ₁₀₀	D ₁₀₀	P ₆₀	E ₁₀₀	I _{82,9}	V _{65,7}	M _{62,9}	H _{94,3}	S _{77,1}	F _{97,1}	N _{94,3}	C ₁₀₀		
	T _{12,9}	N _{11,4}	H ₂₀	A _{2,9}				L _{37,1}		V _{14,3}	T _{28,6}	T _{31,4}	L _{2,9}	T _{17,1}	L _{2,9}	I _{2,9}			
	A _{8,6}	P _{11,4}	P _{7,1}					Q _{2,9}		L _{2,9}	E _{5,7}	A _{2,9}	P _{2,9}	M _{2,9}		T _{2,9}			
	K _{8,6}	R _{11,4}										R _{2,9}		I _{2,9}					
	Q _{2,9}	S _{11,4}																	
	S _{1,4}	K ₁₀																	
		H _{8,6}																	
		E _{1,4}																	

Figura 18: Representação da sequência de aminoácidos consenso do subtipo B correspondente à região C2V3C3 do gene *env* de HIV-1. Por baixo de cada aminoácido da sequência consenso B, representado pelo código universal de uma letra, estão representados os polimorfismos existentes na posição em questão para as sequências estudadas, com as percentagens correspondentes a cada aminoácido (em índice). O símbolo “Δ” representa uma deleção, o símbolo “^” representa a existência de um potencial local de N-glicosilação e o símbolo “*” representa a introdução de um codão *stop* na grelha de leitura. Os números a vermelho correspondem à numeração de aminoácidos no alinhamento da **Figura 17**. A região da ansa V3 (a.a. 91 a 125) está representada a cor azul.

posições de aminoácidos existentes no alinhamento, 119 (69,2%) apresentam pelo menos uma alteração em relação à sequência consenso B, sendo 53 (30,8%) absolutamente conservadas. Estas encontram-se na sua maioria na região C2, 35 posições, e as restantes distribuídas de forma igual entre as regiões V3 e C3, 9 posições conservadas em cada região. Os números encontrados correspondem a uma taxa de conservação absoluta de 38,9%, 25,7% e 19,1%, respectivamente para as regiões estudadas C2, V3 e C3. É nesta última região que se encontram as posições mais variáveis do alinhamento (137 e 141), apresentando frequências para o aminoácido mais representado inferiores a 25%. Entre as posições absolutamente conservadas encontram-se as que correspondem às extremidades da ansa V3 (91C e 125C) e as duas posições da sua coroa (105G e 107G), como se viu anteriormente. Digno também de nota, o facto de nem sempre o aminoácido da sequência consenso B representar o aminoácido mais frequente no alinhamento, reflectindo a eventual presença de “assinaturas aminoacídicas” de subtipos não-B. Para testar esta hipótese, procedeu-se a uma pesquisa sistemática de eventuais substituições associadas, de um modo significativo, a diferentes subtipos, recorrendo a uma análise estatística com os testes exacto de Fisher bilateral ou de χ^2 aplicados a tabelas de contingência 2x2. Os resultados obtidos permitiram identificar algumas “assinaturas aminoacídicas” para os subtipos não-B (T33K, R45K, V64M, F70I, E83K, S84A, R106Q, K129E, I137V, P160L, V163T, M164T), o subtipo A (I100V, A107T, V138A, K139T, K140Q, E143K, S155H) e o subtipo G (V48I, T133M, L134I, V151T), na grande maioria, de um modo altamente significativo ($p < 0,001$). Por fim, a existência de ambiguidades de tradução nas posições 75, 122 e 137 do alinhamento conduziu à introdução de codões de terminação da tradução em algumas das subpopulações virais da *quasispecies*. Estas, no entanto, correspondiam sempre a subpopulações muito minoritárias (0,5-0,9%) (**Figura 18**).

3.5.2. Locais de N-glicosilação

Os locais de N-glicosilação (NxT/S, em que x pode representar qualquer aminoácido, excepto prolina), presentes nas sequências, foram determinados com recurso ao programa *N-GlycoSite*, disponível em <http://www.hiv.lanl.gov/content/sequence/GLYCOSITE/glycosite.html>. Dos 11 locais de N-glicosilação detectados na sequência consenso B (posições 27, 34, 55, 69, 82, 88,

94, 124, 131, 147 e 153) nenhum se encontra totalmente conservado nas 35 sequências estudadas (**Tabela 6**), havendo, no entanto, locais quase completamente conservados (posições 34, 55 e 94). Um número significativo de perdas de locais de N-glicosilação absolutamente conservados está relacionado com a existência de *quasispecies*, nos locais em causa (mantendo-se, em geral, esses locais, numa parte das subpopulações virais da *quasispecies*). Outro fenómeno observado é o da compensação de locais de N-glicosilação perdidos pela selecção de mutações adicionais que tendem a substituir o local perdido por outro, imediatamente a montante (por exemplo, compensação do local 27NxT, em PT_TX18 e PT_TX32) ou a jusante (por exemplo, compensação do local 124NxS, em PT_TP16, PT_TP184 e PT_TP98). No cômputo geral, o número de locais de N-glicosilação encontrados varia entre 6 e 12, com a maioria das sequências a exhibir 10 e 11 locais (10 sequências cada), à imagem da sequência consenso B. Quando se compara a distribuição dos locais de N-glicosilação entre grupos de sequências B e não-B, observa-se uma distribuição genericamente aleatória, exceptuando os casos dos locais 82NxS e 131NxT (**Tabela 6**), para as quais a análise estatística realizada suporta uma associação preferencial com o subtipo B (teste exacto de Fisher bilateral ou teste de χ^2 , $p < 0,01$). Já quanto ao número médio de locais de N-glicosilação nos dois grupos de sequências, este não se mostrou significativamente diferente (teste t de Student, $p = 0,454$), sendo de 9,78 para as sequências B e de 9,42 para as não-B.

Tabela 6: Resultado da pesquisa de locais de N-glicosilação (absolutamente conservados).

	Posição segundo a sequência consenso B										
	27	34	55	69	82	88	94	124	131	147	153
	NxT	NxS	NxS	NxT	NxS	NxT	NxT	NxS	NxT	NxT	NxS
Nº de sequências B (n=23)	20	22	23	19	17	16	22	20	21	22	16
Nº de sequências não-B (n=12)	12	12	11	10	3	9	12	9	5	11	9
Total (n=35)	32	34	34	29	20	25	34	29	26	33	25

[N=asparagina; T=treonina; S=serina; x=qualquer aminoácido, excepto prolina]

3.5.3. Previsão de Utilização de Co-receptores

A previsão do co-receptor utilizado pelas estirpes virais para a entrada nas células hospedeiras foi realizada com recurso a três algoritmos distintos, baseados na estimativa da carga dos resíduos presentes nas posições 11 e 25 da região V3 (57) ou nos programas *FORTINBRAS PSSM* (disponível em <http://fortinbras.us/cgi-bin/fssm/fssm.pl>) e *geno2pheno*_[coreceptor] v.1.2. (implementado em <http://www.viro.med.uni-erlangen.de/>). Adicionalmente, foi ainda realizado o cálculo da carga líquida total da região V3, considerando carga positiva nos resíduos de arginina e lisina e carga negativa nos resíduos de ácido aspártico e ácido glutâmico, uma vez que há dados que sugerem que esta também tem influência no co-receptor utilizado (37). Através da análise da **Tabela 7**, pode verificar-se a existência de resultados discordantes, para algumas sequências.

Através da regra 11/25, baseada na identidade dos aminoácidos presentes nas respectivas posições da ansa V3, foram identificadas 3 (8,6%) amostras (PT_TP107, PT_TX16 e PT_TX32) que utilizam o co-receptor CXCR4. Por sua vez, o programa *PSSM* previu que 11 (31,4%) sequências poderão ser tipo X4 ou apresentar tropismo duplo R5/X4 (PT_TP19, PT_TP156, PT_TX12, PT_TX15, PT_TP107, PT_TP184A, PT_TX16, PT_TX18, PT_TX32, PT_TX45 e PT_TX14). No entanto, este programa tem a desvantagem de utilizar uma matriz baseada exclusivamente em sequências do subtipo B (ou do subtipo C), pelo que a previsão realizada para as amostras de subtipos não-B pode não ser a mais exacta. Finalmente, o programa *geno2pheno* identificou 11 (31,4%) amostras como X4 ou com tropismo duplo R5/X4 (PT_TP02, PT_TP16, PT_TP107, PT_TP114, PT_TP159A, PT_TP184A, PT_TX16, PT_TX30, PT_TX32, PT_TX45 e PT_TP13). Como seria antecipado, os resultados obtidos não são totalmente coincidentes, apesar de se observar uma concordância absoluta no que diz respeito às amostras PT_TP107, PT_TX16 e PT_TX32, que foram identificadas como correspondendo a estirpes X4 ou com tropismo duplo R5/X4 pelos três métodos utilizados. As sequências PT_TP184A e PT_TX45 apresentam concordância dos resultados dos programas *PSSM* e *geno2pheno*, tendo sido também consideradas como X4 ou com tropismo duplo. Todas as outras sequências (n=30) foram consideradas do tipo R5, embora somente 60% destas (n=18) com concordância absoluta de resultados. No cômputo geral, obteve-se uma taxa de concordância entre as previsões efectuadas

com estes dois algoritmos de apenas 65,7% (23/35), sendo as discordâncias mais frequentes nas sequências de subtipos não-B (6/12, i.e., 50,0%) do que nas de subtipo B (6/23, i.e., 26,1%) e nos espécimes com ansas V3 mais curtas (34 aminoácidos) (3/6, i.e., 50,0%) face às mais longas (35 aminoácidos) (9/29, i.e., 31,0%). Estas diferenças, no entanto, não se mostraram estatisticamente significativas (teste exacto de Fisher bilateral, $p=0,261$ e $p=0,391$, respectivamente). A amostra PT_TX16 é a que apresenta a carga líquida total da ansa V3 mais elevada, o que estará associado à utilização do co-receptor CXCR4 (37), corroborando os resultados anteriores. Testou-se ainda a eventual associação entre a natureza da estirpe quanto à utilização do co-receptor (R5 vs. R5X4/X4) e os genótipos virais (B vs. não-B). O teste estatístico realizado comprovou a independência das duas variáveis (teste exacto de Fisher bilateral, $p=0,141$).

Por fim, digno de nota que a substituição I12V (numeração na ansa V3), cuja presença foi apontada por Svicher *et al.* (2011) (150) como significativamente associada ao tropismo duplo R5X4, fenotipicamente definido, foi encontrada em 7 das 35 sequências em estudo (**Figura 17**), havendo, neste trabalho, uma associação evidente da sua presença com as sequências do subtipo A ($p < 0,001$). Já as mutações T2M e I26R, igualmente descritas em associação directa com o tropismo duplo R5X4, não foram encontradas na série em estudo (**Figura 17**).

Tabela 7: Classificação das estirpes virais de acordo com a previsão de utilização de co-receptores realizada através de diferentes algoritmos (com indicação dos aminoácidos das posições 11 e 25 e da carga líquida da ansa V3).

Amostras	a.a. (posições 11/25)	Regra 11/25	PSSM	geno2pheno	Carga líquida ansa V3 (R+K) - (D+E)
TP19	SX(E/D)	R5	R5X4/X4	R5	3
TP156	SS	R5	R5X4/X4	R5	3
TP208	SD	R5	R5	R5	2
TX12	SS	R5	R5X4/X4	R5	3
TX13	SD	R5	R5	R5	3
TX15	SD	R5	R5X4/X4	R5	3
TX41	SD	R5	R5	R5	3
TX46	SD	R5	R5	R5	2
TP02	GE	R5	R5	R5X4/X4	4
TP16	SE	R5	R5	R5X4/X4	5
TP24	GE	R5	R5	R5	3
TP38	SA	R5	R5	R5	4
TP71	SD	R5	R5	R5	3
TP85	SQ	R5	R5	R5	5
TP107	RE	X4	R5X4/X4	R5X4/X4	4
TP114	SE	R5	R5	R5X4/X4	4
TP159A	SD	R5	R5	R5X4/X4	4
TP184A	SD	R5	R5X4/X4	R5X4/X4	3
TP219	SD	R5	R5	R5	3
TX16	RG	X4	R5X4/X4	R5X4/X4	7
TX18	SN	R5	R5X4/X4	R5	5
TX21	SE	R5	R5	R5	3
TX30	X(S/G)X(G/E)	R5	R5	R5X4/X4	4
TX32	RE	X4	R5X4/X4	R5X4/X4	5
TX37	SQ	R5	R5	R5	4
TX40	SG	R5	R5	R5	4
TX45	SQ	R5	R5X4/X4	R5X4/X4	4
TX47	SE	R5	R5	R5	3
TX50	SE	R5	R5	R5	4
TX59	X(G/S)E	R5	R5	R5	3
TX60	SE	R5	R5	R5	4
AS646	SD	R5	R5	R5	4
TP13	SD	R5	R5	R5X4/X4	2
TP98	SA	R5	R5	R5	4
TX14	SS	R5	R5X4/X4	R5	3

[O algoritmo *PSSM* utiliza uma matriz baseada em sequências do subtipo B, pelo que as previsões presentes na tabela para sequências de outros subtipos não são absolutamente fiáveis.

Os aminoácidos são indicados pelo código universal de uma letra; R5 = previsão de utilização exclusiva do co-receptor CCR5; X4 = previsão de utilização exclusiva do co-receptor CXCR4; R5X4 = previsão de utilização de qualquer um dos co-receptores.]

3.5.4. Análise de Polimorfismos Genéticos Associados a Resistência ao Maraviroc

Foi realizada a pesquisa de polimorfismos genéticos naturais associados à resistência ao maraviroc, presentes na região V3 das sequências de HIV-1 estudadas, tendo por base a **Tabela 1**. Esta tabela apresenta uma compilação de mutações associadas à resistência a este fármaco, com origem em dados da literatura (101, 115, 138, 139). Através da análise da **Figura 17** observa-se que a mutação 2I está presente em 3 sequências (PT_TP02, PT_TP159A e PT_TX21), sendo que todas são R5 e do subtipo B (**Tabela 8**). Nesta posição existem ainda 3 sequências (PT_TX30, PT_TX59 e TX_60) que apresentam uma ambiguidade de tradução, existindo para todas a possibilidade de um resíduo de isoleucina ocupar esta posição (**Tabela 5**). A mutação 4L não se apresenta em nenhuma das sequências estudadas, bem como as mutações pontuais 16A, 19S, 21I, as inserções G15_P16INSG/A, N24_I25INSI, a deleção G18DEL e as combinações de mutações 13H/S+26V, 18G+22T, 19S/T+26V, 20F+21I e G15_P16INSG+25D (**Tabela 8**). Para a posição 4, no entanto, podem observar-se 2 sequências (PT_TP19 e PT_TX59) que apresentam uma indeterminação, não se verificando, em nenhum dos casos, a possibilidade da mutação 4L estar presente. Pode observar-se que a mutação 11R está presente em 3 sequências, todas R5X4/X4 e do subtipo B. Nesta mesma posição, a substituição 11S (presente na sequência consenso B) está presente em 28 sequências, quase exclusivamente de tropismo R5 (26/28) e maioritariamente do subtipo B (16/28). Nesta posição existem ainda duas sequências (PT_TX30 e PT_TX59) que apresentam uma ambiguidade de tradução, devida à presença de *quasispecies*, havendo, para ambos os casos, um resíduo de serina numa das subpopulações virais. Quanto à mutação 13H, esta encontra-se presente em 16 sequências, 15 destas R5 e uma R5X4/X4, sendo 11 do subtipo B e 5 de subtipos não-B. Também a mutação 13S está presente em 2 sequências, ambas R5 e de subtipos não-B, nomeadamente, G. Nesta posição existem 3 sequências que apresentam uma indeterminação (PT_TP19, PT_TP219 e PT_TX30). Destas, a primeira pode apresentar a mutação 13H, a última pode apresentar as mutações 13S ou 13H e a sequência PT_TP219 não tem qualquer possibilidade de apresentar a mutação. A mutação 18G está presente numa única sequência (PT_TP24), do subtipo B e de tropismo R5 (nesta posição, a sequência PT_TP219 apresenta uma indeterminação, mas nunca poderá ser

traduzida como um resíduo de glicina). A substituição 19T encontra-se em 6 sequências, todas R5 e de subtipos não-B, mais especificamente, todas de subtipo A. A substituição 20F (presente também na sequência consenso B) encontra-se em 30 sequências, na grande maioria com tropismo R5 (27/30) e de subtipo B (18/30). Quanto à mutação 20L, esta apresenta-se numa única sequência R5, do subtipo B (PT_TP71). Nesta posição, há uma indeterminação de tradução na sequência PT_TP24, ocorrendo a presença quer de um resíduo de leucina, quer de fenilalanina, sendo assim possível a presença de qualquer uma destas mutações. Na posição 22 existem 13 sequências com a substituição 22T, também presente na sequência consenso B, 8 R5, 5 R5X4/X4, 11 do subtipo B e 2 de subtipos não-B. Nesta posição, a sequência PT_TX30 apresenta uma ambiguidade, apresentando a substituição em causa em parte da população da *quasispecies* viral. A mutação 25D está presente em 11 sequências, quase exclusivamente utilizadoras do co-receptor CCR5 (10/11). Entre as 11 sequências, 4 são do subtipo B e 7 de subtipos não-B. A sequência PT_TP19 pode também apresentar esta mutação, sempre que nesta posição for integrado um resíduo de ácido aspártico, que se apresenta como uma das possibilidades de tradução. A mutação 26V encontra-se presente em 2 sequências, uma R5 e outra R5X4/X4, ambas do subtipo B. As amostras PT_TP19, PT_TP24 e PT_TX59 podem também apresentar esta substituição, mas apenas numa parte da *quasispecies*. Entre os padrões mutacionais, apenas dois estão presentes nas sequências em estudo. O padrão 11S+26V, nas sequências PT_TP38 (tropismo R5) e PT_TP184A (tropismo R5X4/X4), ambas do subtipo B. O padrão 20F+25D+26V está presente na sequência PT_TP184A.

A análise de associação efectuada entre a presença de cada uma das diferentes mutações, ou combinações (padrões) de mutações, e a natureza das diferentes sequências, quanto ao genótipo e à utilização de co-receptores, recorrendo, uma vez mais, à construção de tabelas de contingência 2x2, revelou associações estatisticamente significativas apenas entre a presença das mutações 11R ($p=0,002$) e 22T ($p=0,004$) e o tropismo R5X4/X4 e entre a substituição 19T e os genótipos não-B ($p < 0,001$), mais especificamente, o subtipo A.

Tabela 8: Distribuição de mutações relacionadas com resistência ao maraviroc.

Numeração HXB2	Mutações ansa V3	Total de sequências	Sequências R5	Sequências R5X4/X4	Subtipo B	Subtipos não-B
297	2I	3	3	0	3	0
299	4L	0	0	0	0	0
306	11R	3	0	3	3	0
306	11S	28	26	2	16	12
308	13H	16	15	1	11	5
308	13S	2	2	0	0	2
313	16A	0	0	0	0	0
315	18G	1	1	0	1	0
316	19S	0	0	0	0	0
316	19T	6	6	0	0	6
317	20F	30	27	3	18	12
317	20L	1	1	0	1	0
318	21I	0	0	0	0	0
319	22T	13	8	5	11	2
322	25D	11	10	1	4	7
323	26V	2	1	1	2	0
	G15_P16INSG/A	0	0	0	0	0
	N24_I25INSI	0	0	0	0	0
	G18DEL	0	0	0	0	0
Padrões de Mutações						
	11S+26V	2	1	1	2	0
	13H/S+26V	0	0	0	0	0
	18G+22T	0	0	0	0	0
	19S/T+26V	0	0	0	0	0
	20F+25D+26V	1	0	1	1	0
	20F+21I	0	0	0	0	0
	G15_P16INSG+25D	0	0	0	0	0

4. Discussão e Conclusões

4.1. Amplificação do Genoma do HIV-1

A técnica de transcrição reversa *in vitro* seguida de *nested* PCR (RT-*nested* PCR) foi a metodologia adoptada para a amplificação da região C2V3C3 de gp120 do gene *env* do HIV-1 a partir de RNA viral purificado do plasma de 52 amostras de sangue de indivíduos utilizadores de drogas por via endovenosa, seropositivos para HIV-1, da região da Grande Lisboa. Uma vez optimizadas as condições de amplificação, incluindo as sequências dos *primers* a utilizar, a RT-*nested* PCR é um método que garante elevadas sensibilidades e especificidades na amplificação do genoma viral. Para se conseguir obter o produto de amplificação requerido e minimizar o aparecimento de produtos inespecíficos procedeu-se, inicialmente, à optimização das condições experimentais de retrotranscrição e de amplificação, incluindo a escolha dos *primers* que garantem melhores resultados. Durante esta fase de optimização das condições de amplificação, e após a estimativa das temperaturas de desnaturação (T_m) dos diferentes *primers*, testaram-se diferentes temperaturas de hibridação (para a fase de retrotranscrição, 55, 50 e 45°C, e para a 1ª PCR, 55, 50, 45 e 40°C), bem como diversas alternativas de combinação entre eles. Foi também testado o número de ciclos, e a utilização ou não de incrementos de tempo, durante a fase de polimerização da 1ª PCR. A hibridação dos *primers* na sequência alvo do RNA matriz é uma fase fundamental durante a retrotranscrição, assim como a sua subsequente hibridação no DNA complementar obtido após esse primeiro passo experimental. Assim, as temperaturas de hibridação são determinantes para que a ligação dos *primers* ocorra de forma correcta, na sua sequência alvo. Se a temperatura de hibridação for demasiado elevada, pode não ocorrer hibridação, especialmente se a sequência alvo diferir substancialmente da sequência para a qual o *primer* foi desenhado, sendo este facto de especial importância quando se trabalha com genomas que apresentam elevada variabilidade genética, como é o caso do genoma do HIV. Pelo contrário, se a temperatura de hibridação for demasiado baixa, existe um risco aumentado de amplificação de produtos inespecíficos. Este problema ocorre quando a restringência de hibridação se torna muito baixa e os *primers* se ligam em locais diferentes da sequência alvo, sendo que a baixa restringência aumenta também a possibilidade da formação de *primer-dimers* (este fenómeno ocorre quando os *primers* utilizados se ligam inicialmente entre si, através de regiões de

complementaridade parcial nas respectivas extremidades 3', permitindo a amplificação subsequente do dímero formado). Já para a 2ª PCR (*nested*), não se procedeu a qualquer optimização das condições experimentais uma vez que estas já se encontravam optimizadas (42).

Para se evitar a degradação do RNA viral por RNases ambientais e diminuir a possibilidade de ocorrência de contaminação cruzada das amostras durante a sua preparação para amplificação, e tendo em consideração que a técnica de *nested* PCR acarreta esse risco, seguiram-se os seguintes procedimentos:

- 1) realizou-se a preparação das misturas de reacção para PCR numa câmara de fluxo laminar e em espaço físico separado do local onde são adicionadas as amostras de RNA/DNA matriz e onde ocorre a amplificação, dificultando a ocorrência de contaminação com DNA anteriormente amplificado ou outros DNAs contaminantes;

- 2) utilizaram-se pontas de micropipetas com filtro, de forma a reduzir o risco de transferência de ácidos nucleicos contaminantes;

- 3) procedeu-se à descontaminação da superfície de trabalho previamente à adição do RNA matriz nos tubos de PCR;

- 4) utilizou-se o sistema de amplificação *illustraTM puRe Taq Ready-To-Go PCR Beads* (GE Healthcare, Inglaterra) para a segunda reacção de PCR, dado que este sistema implica uma redução substancial do número de pipetagens, reduzindo assim a possibilidade de contaminações, e também da ocorrência de erros, aumentando a reprodutibilidade da técnica.

De realçar ainda a utilização consistente de controlos negativos (ou de contaminação) durante todos os passos do procedimento experimental (retrotranscrição/1ª PCR e 2ª PCR).

Das 52 amostras testadas, foi possível obter o produto de amplificação, com o tamanho esperado aproximado de 565 pb, para 35 (67%). Para as restantes 17 (33%) não foi obtido qualquer produto de amplificação, podendo este facto estar relacionado com vários factores, dos quais se podem destacar, a existência de diferenças entre a sequência dos *primers* e a respectiva região de hibridação e uma baixa quantidade de RNA viral (esta pode ser reduzida porque o indivíduo tinha uma carga viral baixa ou pode ainda ter ocorrido a degradação parcial do RNA, durante a extracção ou até mesmo ainda na amostra de plasma, devido à sua eventual exposição a vários ciclos

prévios de congelação e descongelação). Para 14/35 (40%) dos produtos amplificados, foi necessário recorrer-se a um passo intermédio de clonagem molecular, antes da purificação para sequenciação nucleotídica, dada a quantidade reduzida do material amplificado e/ou o seu acompanhamento com alguns produtos inespecíficos de amplificação.

Por fim, observou-se a inexistência de qualquer influência das características epidemiológicas ou clínicas conhecidas dos indivíduos participantes neste estudo directamente no sucesso da amplificação, com uma única excepção, relativa à categoria “terapia anti-retroviral”. Neste caso, o teste estatístico realizado suporta um muito maior sucesso de amplificação nas amostras provenientes de indivíduos sem terapia. Este facto poder-se-ia explicar, de um modo indirecto, assumindo que indivíduos sem terapia têm, em geral, cargas virais mais elevadas, facilitadoras, portanto, da amplificação. No entanto, este efeito não foi observado quando se testou a associação directa entre o sucesso/insucesso de amplificação e os diferentes níveis de carga viral (conhecidos, no entanto, apenas para 21 dos 52 indivíduos).

4.2. Análise das Sequências Nucleotídicas

Os amplicões obtidos foram enviados para sequenciação nucleotídica em ambas as cadeias, procurando assim minimizar-se a eventual introdução de erros inerentes à própria técnica de sequenciação. As sequências nucleotídicas obtidas das cadeias simples de cada amostra amplificada foram comparadas entre si e editadas. Decidiu-se clonar os amplicões cujas sequências apresentavam uma variabilidade elevada, o que impossibilitava a sua análise correcta, devendo-se este facto à existência frequente de *quasispecies* complexas, tal como seria antecipado considerando a região do genoma viral em estudo. No entanto, tentou minimizar-se a utilização da clonagem, nesta fase, pois, para além de mais dispendiosa e demorada, a sua utilização leva à perda de alguma informação, possivelmente relevante para o estudo, uma vez que a variabilidade verificada nas sequências se deve, como se viu, à existência de mais de uma população viral em circulação, no indivíduo infectado, ou seja populações virais com genomas muito semelhantes entre si, mas com alterações pontuais nas suas sequências nucleotídicas. As *quasispecies* são um resultado da elevada taxa de evolução do HIV-1, que ocorre como consequência das elevadas taxas de mutação e replicação do vírus, da

recombinação genética e como resposta a pressões exercidas pelo sistema imunitário do hospedeiro e da própria terapia anti-retroviral (94, 126, 154). Assim, nesta fase, optou-se pela clonagem apenas das amostras que, após sequenciação, apresentavam cromatogramas de muito difícil interpretação e que após re-sequenciação continuavam a apresentar os mesmos problemas.

Após a edição das sequências, procedeu-se a uma pesquisa preliminar de homologia genética utilizando o programa *BLAST*, com o objectivo de se confirmar rapidamente a identidade das sequências obtidas, sendo também determinados os respectivos subtipos, por análise filogenética automatizada, utilizando-se para esse fim o programa *REGA HIV-1 & 2 Automated Subtyping Tool v.2.0*. De forma a confirmar os resultados obtidos, procedeu-se à construção de uma árvore filogenética com as sequências nucleotídicas em estudo e sequências de referência dos diferentes subtipos, sub-subtipos e duas formas recombinantes circulantes (CRF) de HIV-1 (grupo M).

O método matemático utilizado para a construção da árvore filogenética foi o de *neighbor-joining* (135). Este é um método largamente utilizado para a construção de árvores filogenéticas, pois é rápido e produz uma única árvore como resultado (118). O método de substituição de nucleótidos, utilizado na construção da árvore para o cálculo da matriz de distâncias genéticas entre as diferentes sequências em estudo, foi o modelo de 2 parâmetros de Kimura (91), o qual distingue dois tipos de substituições nucleotídicas: transversões (substituição de uma purina por uma pirimidina ou vice-versa) e transições (purina-purina ou pirimidina-pirimidina). Este modelo assume que as transições tendem a ocorrer mais frequentemente que as transversões (78). Para avaliar a robustez da inferência filogenética obtida da árvore filogenética construída, foi utilizado o teste de *bootstrapping*. Este método recorre a partes aleatórias do alinhamento para realizar uma série de árvores filogenéticas individuais e avaliar em quantas destas está presente um determinado agrupamento de sequências, obtendo-se assim uma percentagem que indica o número de árvores em que aquele agrupamento é mantido. O teste de *bootstrapping* permite assim avaliar se um determinado agrupamento ocorre por similaridade filogenética real das sequências emparelhadas ou simplesmente de uma forma aleatória (7).

A análise da árvore filogenética (**Figura 14**) revela uma distribuição desigual das amostras em estudo entre 4 subtipos (B, A, G e F1), sendo o subtipo B claramente

predominante (23/35, 65,7%), corroborando resultados anteriores de Esteves *et al.* (2002) (48) e Duque *et al.* (2003) (45). Contudo, e em semelhança ao apresentado também por estes autores, a presença de outros subtipos e CRFs demonstra um padrão epidemiológico distinto do da maioria dos países europeus e norte-americanos. A origem de subtipos não-B em Portugal pode dever-se à imigração de indivíduos de origem africana e às deslocções de indivíduos portugueses de e para países africanos (48). Nos últimos anos, têm ocorrido a Portugal indivíduos de outras origens geográficas (Europa de Leste, América do Sul), o que poderá ainda contribuir para uma maior diversificação de subtipos.

Como se pode verificar na **Figura 14**, as sequências do grupo B (n=23) apresentam uma elevada diversidade, encontrando-se distribuídas por diferentes subgrupos, demonstrando uma elevada variedade de formas virais do subtipo B em circulação em Portugal. Um destes subgrupos emparelha duas sequências em estudo (PT_TX32 e PT_TX45) com sequências de referência da forma recombinante circulante CRF14_BG, originárias de Espanha (a sequência X605 é proveniente de um indivíduo toxicodependente de origem espanhola e a sequência X623 tem origem num indivíduo homossexual português mas residente em Espanha) (40). Este subgrupo apresenta um alto valor de *bootstrap* (99%) indicando uma elevada consistência das relações filogenéticas entre estas sequências. Estes resultados indiciam que as sequências recombinantes 14_BG dos vírus circulantes em toxicodependentes portugueses e espanhóis são semelhantes do ponto de vista filogenético, comprovando uma origem evolutiva comum (40). Assim, pode inferir-se que as sequências PT_TX32 e PT_TX45 pertencem à forma recombinante circulante CRF14_BG ou que estão filogeneticamente próximas das estirpes B suas parentais. Só o conhecimento de uma porção significativa do genoma de ambas as estirpes permitiria ajudar a distinguir entre estas duas hipóteses. As restantes sequências do subtipo B encontram-se distribuídas por mais cinco subgrupos, mas com valores de *bootstrap* não significativos, o que demonstra uma baixa consistência das relações filogenéticas entre as sequências presentes nestes grupos, resultado da elevada diversidade genética apresentada.

As sequências genotipadas como F1 (n=1) e G (n=3) encontram-se agrupadas com as respectivas sequências de referência do sub-subtipo F1 e do subtipo G, formando agrupamentos filogeneticamente muito robustos (**Figura 14**). Estes resultados

atestam uma elevada consistência da relação filogenética estabelecida nestes grupos, permitindo assim confirmar o resultado da subtipagem automatizada realizada previamente. Como curiosidade, de referir que embora se encontrem emparelhadas com uma sequência de referência de Portugal e outra de Espanha, as sequências G em estudo apresentam maior proximidade filogenética com a primeira, como seria antecipado. A sequência de referência de Espanha ramifica separadamente das restantes, resultando assim um grupo composto pela sequência de referência de Portugal e as três amostras em estudo do subtipo G. No entanto, e apesar de algo divergente das sequências de origem portuguesa, a sequência de referência de Espanha partilha claramente um ancestral comum com aquelas, corroborando dados anteriores sobre a circulação do HIV-1 do subtipo G na Península Ibérica (40). Um padrão mais homogêneo poderia eventualmente ser observado se se tivesse introduzido um maior número de sequências de referência deste subtipo para a construção da árvore filogenética.

Relativamente ao agrupamento das sequências do subtipo A, este é o menos consistente de todos (**Figura 14**), eventualmente porque se optou pela introdução de referências representantes de todos os sub-subtipos do subtipo A e também da CRF02_AG (a região C2V3C3 do gene *env* é do subtipo A nesta forma recombinante circulante). Este agrupamento engloba 8 sequências portuguesas em estudo, parecendo haver a tendência de formação de um subgrupo individualizado, embora sem consistência filogenética evidente (valor de *bootstrap* não significativo). Tal poderá dever-se à circulação em Portugal de formas virais divergentes do subtipo A, pelo menos relativamente às sequências de referência utilizadas na construção da árvore filogenética. Há uma sequência (PT_TX46) que se encontra agrupada com uma referência CRF02_AG de Espanha, mas com um *bootstrap* inferior a 70%, o que revela uma baixa consistência das relações filogenéticas. Este grupo partilha um ancestral comum com o conjunto que contém todas as restantes sequências do subtipo A em estudo. Neste grupo, a mencionar, duas sequências (PT_TX12 e PT_TP156) que se encontram emparelhadas com um *bootstrap* de 100% e suportadas por ramos muito curtos. Estes resultados indicam que estas sequências apresentam uma divergência genética muito baixa entre si. Isto pode ter resultado da infecção de ambos os indivíduos por uma mesma estirpe viral, através da partilha de seringas, algum tempo antes da colheita das amostras, o que permitiu, ainda assim, o aparecimento de alguma

divergência (especialmente no caso da região em estudo, pois é uma das mais variáveis do genoma viral, sendo influenciada pela pressão selectiva exercida pelo sistema imunitário do hospedeiro, o que leva a uma rápida evolução divergente).

Como referido anteriormente, o subgrupo que contém as sequências PT_TX46 e a sequência de referência CRF02_AG de Espanha apresenta um valor de *bootstrap* inferior a 70%, não sendo assim possível inferir a partir deste resultado se a sequência PT_TX46 pertence de facto à CRF02_AG. Assim, realizou-se uma análise de *bootscanning*, para confirmar este resultado, com recurso ao programa *SimPlot* (105). Este programa permite inferir sobre a existência de pontos de recombinação em determinada sequência. A análise é feita a partir do alinhamento da sequência em estudo com um conjunto de sequências de referência. O alinhamento é dividido em segmentos sobreponíveis que se vão deslocando ao longo do seu comprimento. Cada segmento do alinhamento é utilizado para construir a árvore filogenética respectiva, com indicação gráfica dos valores de *bootstrap* relativos aos agrupamentos que suportam a sequência em análise com a referência filogeneticamente mais próxima (105). Assim, de um modo indirecto, é possível identificar os subtipos envolvidos numa eventual recombinação e o local onde aquela terá ocorrido. Os resultados obtidos não permitiram chegar a uma conclusão definitiva, pelo que a sequência PT_TX46, face aos dados disponíveis, foi simplesmente considerada como do subtipo A.

Com este resultado foi possível estabelecer a distribuição dos subtipos de HIV-1 entre as amostras estudadas (**Figura 15**), confirmando o que havia sido anteriormente determinado pelo programa *REGA HIV-1 & 2 Automated Subtyping Tool v.2.0*, uma vez que as duas sequências agora classificadas como CRF14_BG e a sequência eventualmente CRF02_AG correspondem, na região *env* estudada, aos subtipos B e A, respectivamente.

Uma segunda árvore filogenética (**Figura 16**) foi construída a partir de um alinhamento múltiplo das sequências em estudo com sequências de referência e sequências retiradas da base de dados referentes a amostras de Portugal, recorrendo aos métodos descritos anteriormente. Esta árvore permitiu analisar as relações filogenéticas estabelecidas entre as sequências em estudo e outras previamente estudadas de origem portuguesa. A topologia obtida confirmou, por um lado, as relações filogenéticas inferidas na árvore da **Figura 14** e, por outro lado, a ausência de qualquer tipo de

segregação, nomeadamente, de tipo cronológico, ou de origem (conhecida) das amostras, entre as sequências portuguesas previamente estudadas e as novas sequências, havendo intercalação entre ambos os tipos e entre aquelas e as referências utilizadas. Da análise desta árvore, foi possível observar que as sequências em estudo estabelecem relações filogenéticas próximas com outras anteriormente descritas em Portugal, contudo os baixos valores de *bootstrap*, em alguns grupos, não permitiram a confirmação inequívoca destas relações filogenéticas. Pelo mesmo motivo, não foi possível tornar mais consistentes alguns agrupamentos, pouco consistentes, da árvore filogenética da **Figura 14**, nomeadamente, os referentes ao subtipo A.

4.3. Análise das Sequências de Aminoácidos

Através da análise das sequências de aminoácidos (**Figuras 17 e 18**) correspondentes à região C2V3C3 da gp120 do gene *env* das estirpes de HIV-1 em estudo, pode verificar-se que a região C2 é a mais conservada das regiões estudadas. A região V3, apesar de ser das mais variáveis do genoma, apresenta posições absolutamente conservadas (n=9), como é o caso das cisteínas nas posições 91 e 125 responsáveis pela formação da ponte dissulfeto que mantém a ansa V3 (18). A conservação absoluta desses resíduos de aminoácidos reflectirá a sua importância na manutenção da estrutura terciária desta região da gp120, absolutamente necessária, por sua vez, para a manutenção da funcionalidade da própria glicoproteína. Relativamente aos motivos da coroa da ansa V3, foram encontrados cinco motivos distintos, com a sequência GP/LGG/K/Q/R. Os motivos GPGR e GPGQ foram os mais frequentemente encontrados (no seu conjunto, em 27 sequências, sendo GPGR o mais frequente), tendo o primeiro sido associado, de um modo significativo, ao subtipo B e o segundo a subtipos não-B, mais propriamente, aos subtipos A e G. Resultados semelhantes foram anteriormente também descritos por outros autores (2, 49). O motivo GLGR apresentado por apenas 2 sequências, do subtipo B, é também descrito por Alencar *et al.* (2010) (2) como estando presente em estirpes do subtipo B tipicamente brasileiro. Estas estirpes do subtipo B característico do Brasil apresentam um resíduo de triptofano ou um dos seus derivados metabólicos (R, V, G, L, E, F, A e S) na coroa da ansa V3, como se pode observar nas duas sequências em estudo. Como se viu, as sequências PT_TP219 e PT_TX45 apresentam uma ambiguidade de tradução nesta região, podendo coexistir

subpopulações virais com os motivos GPGR ou GPGK e GPGK ou GTGK, respectivamente. Enquanto os primeiros foram encontrados noutras sequências, o último motivo possível para a sequência PT_TX45 (GTGK) é único neste conjunto de sequências em estudo.

A região V3 é de especial importância devido ao seu envolvimento no tipo de co-receptor utilizado e no tropismo celular. Esta região sofre mutações a um ritmo acelerado para escapar a pressões imunológicas, nomeadamente relativas ao braço humoral da resposta imunitária do hospedeiro, mas certos segmentos da ansa V3 mantêm-se mais conservados do que outros. Estes últimos são provavelmente essenciais para o vírus. Assim, o vírus altera a ansa V3 apenas o suficiente para escapar a pressões imunológicas, sem perder, contudo, a sua capacidade de reconhecer o co-receptor celular, necessário à entrada do vírus na célula, e de replicar no hospedeiro. As regiões N-terminal (91-99), C-terminal (117-125) e a coroa (105-108) da ansa V3 são mais conservadas que as regiões que flanqueiam a coroa (18), sendo este mesmo resultado observado nas amostras estudadas.

Por fim, a pesquisa sistemática da presença de substituições aminoacídicas associadas, de um modo significativo, aos diferentes subtipos, permitiu identificar algumas “assinaturas aminoacídicas” para os subtipos não-B, na generalidade (n=12), o subtipo A (n=7) e o subtipo G (n=4).

4.4. Locais de N-glicosilação

A extensa N-glicosilação da gp120 do HIV-1 é essencial para que a proteína adquira a sua estrutura correcta (97) e permite ainda ao vírus proteger-se da acção inibitória exercida por anticorpos, através da formação de uma interface altamente glicosilada que foi designada por “face imunologicamente silenciosa” (do inglês *immunologically silent face*) (114). Wei *et al.* (2003) (164) atribuíram um comportamento de sinergismo entre os locais de N-glicosilação da glicoproteína para a formação de um “escudo de glicano” que impede o acesso de anticorpos neutralizantes, permitindo o reconhecimento de novas células susceptíveis para a entrada e replicação.

Foram encontrados 11 possíveis locais de N-glicosilação na sequência consenso B, nas posições 27, 34, 55, 69, 82, 88, 94, 124, 131, 147 e 153 (**Tabela 6**). Os locais encontrados correspondem aos descritos por outros autores (99, 100) na sequência de

referência do subtipo B HXB2, com exceção do local na posição 153 [no entanto, este local, assim como os restantes, foram descritos por Esteves *et al.* (2002) (48)]. Destes locais, 6 encontram-se na região C2, um na ansa V3 e 4 na região C3 e nenhum deles se encontra absolutamente conservado, sendo que os mais conservados (34, 55 e 94) se encontram presentes em 34 sequências. O possível local de N-glicosilação presente na ansa V3 (94NxT) encontra-se entre os mais conservados, enquanto os locais que a flanqueiam (82NxS, 88NxT e 124NxS), pelo contrário, se encontram entre os mais variáveis. Os restantes locais descritos para a região C2 (27NxT, 34NxS e 55NxS) correspondem aos mais conservados, indo de encontro ao descrito por Esteves *et al.* (2002) (48).

Além dos locais descritos anteriormente relativos à sequência consenso B, foram encontrados outros que não estão presentes nesta sequência. Estes locais suplementares foram encontrados, maioritariamente, em sequências que perderam pelo menos um dos locais característicos do consenso B. A sequência PT_TP71 é a única que apresenta um local extra, na posição 23, sem que exista perda de um dos outros locais. As sequências PT_TX14, PT_TP98 e PT_TP159A apresentam dois destes locais suplementares e têm ainda em comum a perda de três dos locais descritos anteriormente, dois dos quais são os mesmos para as três sequências (posições 88 e 131). Assim, em termos genéricos, parece haver um mecanismo de selecção de locais de N-glicosilação alternativos, sempre que, por introdução de mutações, se assiste à perda de locais tipicamente presentes (fenómeno de compensação). Este facto parece, mais uma vez, colocar em evidência a importância para o vírus da manutenção do referido “escudo de glicano”, fundamental para a manutenção da sua estratégia de escape imunológico.

No que se refere a eventuais diferenças entre sequências, segundo o seu subtipo (B *vs.* não-B), assiste-se a uma distribuição quase aleatória dos locais possíveis de N-glicosilação, exceptuando os casos dos locais 82NxS e 131NxT (**Tabela 6**), cuja conservação surge preferencialmente associada ao subtipo B. No entanto, não é claro se este facto terá alguma relevância biológica, já que o número médio de locais de N-glicosilação nos dois grupos de sequências não se mostrou significativamente diferente.

4.5. Previsão da Utilização de Co-receptor

A utilização de fármacos antagonistas do co-receptor CCR5 é limitada a indivíduos infectados nos quais só são detectados vírus com tropismo R5, pelo que a determinação do tropismo viral é requerida antes da implementação de qualquer tratamento com maraviroc. Por esta razão, têm sido desenvolvidos vários ensaios para a previsão da utilização de co-receptores, sendo o ensaio *Trofile*[®] (Monogram Biosciences, South San Francisco, EUA), um ensaio fenotípico, o mais utilizado até à data. Estes testes, apesar de fiáveis, são muito dispendiosos, morosos e requerem instalações especializadas (124). Em alternativa, o co-receptor utilizado pelo vírus pode ser previsto por métodos mais simples, com recurso à determinação da sequência aminoacídica da região V3 da gp120, uma vez que esta região é o principal determinante do tropismo viral (79).

A natureza da sequência aminoacídica da ansa V3 pode ser analisada através da regra 11/25. Esta regra baseia-se no facto da introdução de aminoácidos com carga positiva (arginina, lisina) nestas posições da ansa V3 se encontrar associada a um fenótipo indutor de sincícios e a uma utilização do co-receptor CXCR4 para a entrada nas células do hospedeiro (57). A carga líquida total da ansa V3 pode também ser usada para a previsão do co-receptor utilizado, sendo que uma carga elevada se encontra relacionada com a utilização do co-receptor CXCR4 (37, 57, 129). Neste trabalho, este valor foi calculado subtraindo o número de aminoácidos com carga negativa (ácido aspártico e ácido glutâmico) ao número de aminoácidos com carga positiva (arginina e lisina). A utilização da regra 11/25 em conjunto com a regra da carga líquida total da ansa V3 permite a obtenção de resultados mais fiáveis do que a sua utilização em separado (129). Segundo Raymond *et al.* (2008) (129), a previsão da utilização de CXCR4 como co-receptor, utilizando a regra 11/25 e o valor da carga da ansa V3, em conjunto, requer pelo menos um dos requisitos listados: (i) R ou K na posição 11 da ansa V3 e/ou um K na posição 25; (ii) R na posição 25 e uma carga líquida total da ansa V3 igual ou superior a 5; ou (iii) uma carga líquida total da ansa V3 igual ou superior a 6. A utilização destes critérios permitiu a identificação de apenas três sequências denotando tropismo X4: PT_TP107, PT_TX16 e PT_TX32 (**Tabela 7**).

Para além dos métodos anteriormente descritos, várias ferramentas bioinformáticas têm sido desenvolvidas para a previsão mais fiável da utilização do co-

receptor através da análise da sequência de aminoácidos da ansa V3, como são exemplo o programa *PSSM* e *geno2pheno*_[coreceptor], entre outros (122). O primeiro é um método bioinformático simples que estima a propensão de sequências aminoacídicas da ansa V3 utilizarem o co-receptor CXCR4. O *PSSM* utiliza matrizes, de valores de probabilidade que são atribuídos para cada posição da sequência, construídas a partir de um conjunto de sequências V3 de estirpes virais com fenótipo conhecido. Nestas matrizes, as colunas representam posições de aminoácidos na ansa V3 e as linhas representam os possíveis resíduos aminoacídicos. A cada sequência é atribuída uma pontuação somando os valores das células da matriz que correspondem ao resíduo presente na sequência em cada posição da ansa V3. Estas pontuações são comparadas com os valores-limite estabelecidos originalmente para tropismo X4 (ou duplo R5X4) e R5 (84, 124). Com o intuito de melhorar o grau de certeza com que a previsão da utilização do co-receptor é realizada, Poveda *et al.* (2009) (124) descreveram uma alteração ao algoritmo, baseada no estudo de uma série considerável de estirpes virais com resultados disponíveis de fenotipagem. Uma das grandes desvantagens deste programa reside no facto das matrizes disponíveis serem baseadas exclusivamente em sequências de ansas V3 de vírus de subtipos B ou C, não sendo eventualmente totalmente adequadas aos outros subtipos não-B. Garrido *et al.* (2008) (65) demonstraram que a utilização da matriz B do programa *PSSM* para a previsão do co-receptor de sequências não-B estava associada uma sensibilidade e especificidade menores, em comparação com as sequências B. Neste estudo foi utilizada uma matriz baseada em sequências do subtipo B para as quais o co-receptor utilizado foi previamente determinado (matriz B X4/R5, Poveda2009) (84, 124).

O programa *geno2pheno* utiliza um método de previsão diferente do programa *PSSM*, contudo, em semelhança ao anterior, a previsão é também baseada num conjunto de sequências de diferentes subtipos que foram utilizadas para o seu desenvolvimento. Apesar de outros subtipos estarem presentes neste conjunto de sequências utilizadas para a geração do programa, o subtipo B era o mais representado (143). Neste programa, além da sequência, propriamente dita, pode ainda ser utilizada informação adicional, clínica e estrutural (quando disponível), para uma previsão melhorada (143). Um importante ponto a ter em conta na utilização do programa *geno2pheno* é a determinação da taxa de falsos positivos. A fixação de uma elevada taxa de falsos

positivos possibilita a detecção de um maior número de estirpes X4, mas com uma menor especificidade. Assim, é importante encontrar um equilíbrio entre a sensibilidade e especificidade da previsão (130). Neste estudo foi utilizada uma taxa de falsos positivos de 20%, de acordo com as directrizes europeias (à altura que se iniciou a redacção desta tese) sobre a gestão clínica dos testes de tropismo de HIV-1 (161).

Através da análise da **Tabela 7** pode verificar-se que os diferentes métodos utilizados, baseados em sequências aminoacídicas, nem sempre forneceram resultados concordantes, o que comprova a necessidade destes sofrerem ainda aperfeiçoamentos, dada a sua utilidade na gestão clínica dos indivíduos infectados (são mais rápidos e muito mais baratos do que os testes fenotípicos). O programa *PSSM* previu que 11 amostras terão tropismo X4, ou duplo, o mesmo número do programa *geno2pheno* (mas não necessariamente as mesmas amostras). A análise das sequências recorrendo à regra 11/25, conjugada com a da carga líquida total da ansa V3, permitiu prever a utilização do co-receptor CXCR4 apenas por 3 sequências (sendo estas coincidentes com os resultados dos programas anteriores). Assim, consideraram-se com tropismo X4, ou R5X4, todas as sequências para as quais se obtiveram resultados coincidentes pelo menos com dois dos algoritmos utilizados. Deste modo, consideraram-se 5 sequências nestas condições: PT_TP107, PT_TX16, PT_TX32, PT_TP184A e PT_TX45. De realçar que as três primeiras obtiveram o mesmo resultado nos três métodos utilizados para a previsão. No cômputo geral, as discordâncias de previsão foram mais frequentes nas sequências de subtipos não-B e nos espécimes com ansas V3 mais curtas (34 aminoácidos), o que poderá reflectir uma menor adequabilidade das matrizes utilizadas para estas circunstâncias.

Os diferentes resultados aqui encontrados podem dever-se às diferentes sensibilidades e especificidades apresentadas pelos métodos utilizados para a previsão. Os programas seleccionados neste trabalho encontram-se entre os que apresentam maior sensibilidade, mas com valores a rondar os 60% (130). Poveda *et al.* (2009) (124) realizaram uma comparação entre diferentes métodos bioinformáticos e o ensaio fenotípico *Trofile*[®] (Monogram Biosciences, South San Francisco, EUA), revelando uma concordância de resultados de 80% para o programa *PSSM*, com a matriz X4R5 para sequências B, de 75% para a regra 11/25 e de apenas 66% para o programa *geno2pheno*, com a taxa de falsos positivos de 20%. Estes resultados demonstraram

que, apesar das limitações existentes, estes algoritmos/programas bioinformáticos podem ser úteis para avaliar a possível utilização de fármacos antagonistas do co-receptor CCR5 em locais onde os ensaios fenotípicos não constituem uma opção (123).

4.6. Pesquisa de Polimorfismos Genéticos Associados a Resistência ao Maraviroc

A resistência ao maraviroc pode ocorrer de duas formas, sendo que uma delas é através de substituições de aminoácidos, dependentes da ocorrência de mutações na respectiva região codificante, na região variável V3 da gp120, noutra região da gp120 ou na gp41 (168). Este estudo centrou-se na pesquisa de mutações associadas a resistência presentes na região V3 da gp120. Para tal foi realizada uma pesquisa nas sequências aminoacídicas traduzidas (**Figura 17**) das diferentes mutações descritas na literatura, compiladas na **Tabela 8**. Verificou-se a presença de várias mutações previamente associadas à diminuição da susceptibilidade ao maraviroc nas sequências em estudo, apesar das amostras sanguíneas terem tido origem em indivíduos nunca submetidos a tratamento com este anti-retroviral. A presença frequente destas mutações terá a ver com o facto da região em estudo ser uma das mais variáveis do genoma do HIV-1. Assim, é possível o aparecimento destas mutações em pacientes sem historial de tratamento prévio com maraviroc, bem como o aparecimento de diferentes susceptibilidades ao fármaco antes do início do tratamento (128).

A mutação mais frequente foi a 20F, presente em 30 sequências (85,7%). Este resultado está de acordo com o descrito por Saracino *et al.* (2010) (138) e Seclén *et al.* (2010) (139), para os quais esta mutação foi também a mais observada. A mutação 20F foi mais frequente em sequências R5 (27/30, 90%) do que em sequências R5X4/X4 (3/5, 60%), o que corresponde ao descrito por Saracino *et al.* (2010) (138), mas tal diferença não demonstrou ser significativa. A segunda substituição mais frequente foi a 11S, presente em 28 sequências (80%), o que corresponde a uma frequência muito próxima da descrita por Saracino *et al.* (2010) (138) que reportaram a sua presença em 85% de 1200 sequências analisadas. Nesse mesmo estudo, estes autores descreveram outra substituição muito frequente (25D, em 45% das sequências estudadas), a qual foi também encontrada em cerca de 1/3 das sequências estudadas neste trabalho. Outras mutações muito frequentes foram 13H e 22T, esta última presente em associação significativa com as sequências de tropismo R5X4/X4. As mutações 19T e 26V,

associadas com a resistência *in vitro* ao maraviroc (168), foram observadas em 6 e 2 sequências, respectivamente, sendo que a segunda pode ainda estar presente em 3 outras sequências (PT_TP19, PT_TP24 e PT_TX59) dependendo do aminoácido integrado (ambiguidade de tradução). A estirpe PT_TP19 apresenta a substituição 19T e, numa das suas subpopulações virais, a 26V. Assim, esta amostra pode apresentar ambas as mutações, facto que pode ser importante relativamente à possibilidade de tratamento futuro com este fármaco do indivíduo infectado. A mutação 19T surgiu unicamente em sequências de subtipo não-B, havendo uma associação estatisticamente significativa com o subtipo A, o que vai de encontro ao descrito por Saracino *et al.* (2010) (138) que encontraram esta mutação maioritariamente em sequências de subtipo não-B. As substituições 4L e 19S não foram encontradas e a mutação 11R está presente em 3 sequências, todas com tropismo R5X4/X4 (associação significativa) e do subtipo B. Estas três mutações foram associadas, no ensaio MOTIVATE, com uma baixa taxa de resposta virológica ao maraviroc (101). Os resultados agora obtidos são, no entanto, contrários aos reportados por Seclén *et al.* (2010) (139), os quais encontraram as mutações 4L e 19S nas sequências em estudo, com ausência da mutação 11R, embora, em ambos os casos, as mutações 4L e 19S se apresentassem num baixo número de sequências. O mesmo ocorreu para a substituição 21I, que também não foi encontrada neste trabalho. Outras alterações, como a mutação 16A, as inserções (G15_P16INSG/A e N24_I25INSI) ou a deleção (G18DEL) não foram detectadas em nenhuma das sequências em estudo, o que vai de encontro ao verificado por Seclén *et al.* (2010) (139). Outras mutações associadas na literatura a uma diminuição da susceptibilidade ao maraviroc (2I, 13S, 18G, 20L) foram raramente encontradas na série de sequências aqui em estudo. Entre os vários padrões mutacionais analisados, apenas dois foram encontrados: 20F+25D+26V, numa sequência, e 11S+26V, em duas sequências. Este resultado corrobora os resultados de vários autores (138, 139, 146), nos quais estes padrões mutacionais foram encontrados entre as sequências em análise, sendo o último padrão também o mais frequente.

Em conclusão, foi possível constatar um elevado número de vírus do subtipo B a circular entre toxicodependentes por via endovenosa na área da Grande Lisboa, mas também a presença de vírus de outros subtipos (A, G, F1). Este resultado vai de

encontro ao descrito em estudos anteriores que estabeleceram um aumento da circulação de vírus de subtipos não-B em Portugal (45, 48, 49). Verificou-se também uma elevada diversidade genética entre as sequências em estudo e entre estas e sequências anteriormente descritas para o nosso país, especialmente para os subtipo B (incluindo a CRF14_BG) e A (incluindo, eventualmente, a CRF02_AG), o que reflectirá uma elevada diversidade, mesmo intragenotípica, dos vírus circulantes em Portugal. Constatou-se igualmente que os métodos genotípicos para a previsão da utilização de co-receptores virais apresentam ainda limitações, principalmente ao nível da previsão em estirpes de subtipos não-B, revelando-se este facto de especial importância considerando a alteração do padrão de distribuição geográfica destas estirpes, fundamentalmente nos países do hemisfério Norte. Apesar de algumas limitações, os métodos genotípicos podem, ainda assim, ser considerados como alternativas viáveis sempre que a utilização dos ensaios fenotípicos não é possível, dada a sua complexidade e custos muito elevados. Finalmente, observou-se a presença de várias mutações e de alguns padrões mutacionais associados à resistência ao maraviroc nas sequências em estudo. Uma vez que estas sequências são provenientes de vírus que infectam indivíduos que nunca foram submetidos a tratamento com este fármaco, a presença destas mutações resultará da elevada variabilidade da região em estudo, em resposta à pressão exercida pelo sistema imunitário do hospedeiro. O impacto da sua presença é ainda desconhecido, sendo necessário o desenvolvimento de estudos mais aprofundados de modo a perceber quais os efeitos da presença destas mutações antes do início do tratamento, considerando até a sua rápida transmissão neste tipo de população, através da partilha de materiais contaminados.

5. Referências Bibliográficas

1. **Abram, M. E., A. L. Ferris, W. Shao, W. G. Alvord, and S. H. Hughes.** 2010. Nature, Position, and Frequency of Mutations Made in a Single Cycle of HIV-1 Replication. *J. Virol.* **84**: 9864-9878.
2. **Alencar, C. S., A. S. Nishiya, S. Ferreira, M. T. M. Giret, R. S. Diaz, and E. C. Sabino.** 2010. Evaluation of Primary Resistance to HIV Entry Inhibitors Among Brazilian Patients Failing Reverse Transcriptase/ Protease Inhibitors Treatment Reveal High Prevalence of Maraviroc Resistance-Related Mutations. *AIDS Res. Hum. Retroviruses.* **26**: 1267-1271.
3. **Alkhatib, G., C. Combadiere, C. C. Broder, Y. Feng, P. E. Kennedy, P. M. Murphy, and E. A. Berger.** 1996. CC CKR5: A RANTES, MIP-1 α , MIP-1 β Receptor as a Fusion Cofactor for Macrophage-Tropic HIV-1. *Science.* **272**: 1955-1958.
4. **Allan, J. S., J. E. Coligan, T. H. Lee, M. F. McLane, P. J. Kanki, J. E. Groopman, and M. Essex.** 1985. A new HTLV-III/LAV encoded antigen detected by antibodies from AIDS patients. *Science.* **230**: 810- 813.
5. **Arenzana-Seisdedos, F.** 2008. La entrada viral como diana terapêutica. Situación actual de los inhibidores de la entrada. *Enferm. Infecc. Microbiol. Clin.* **26**: Suppl 11: 5-11.
6. **Arya, S. K., C. Guo, S. F. Josephs, and F. Wong-Staal.** 1985. Trans-activator gene of human T-lymphotropic virus type III (HTLV-III). *Science.* **229**: 69-73.
7. **Baldauf, S. L.** 2003. Phylogeny for the faint of heart: a tutorial. *Trends Genet.* **19**: 345-351.
8. **Bártolo, I., C. Rocha, J. Bartolomeu, A. Gama, R. Marcelino, M. Fonseca, A. Mendes, M. Epalanga, P. C. Silva, and N. Taveira.** 2009. Highly divergent subtypes and new recombinant forms prevail in the HIV/AIDS epidemic in Angola: New insights into the origins of the AIDS pandemic. *Infect. Genet. Evol.* **9**: 672-682.
9. **Bártolo, I., A. B. Abecasis, P. Borrego, H. Barroso, F. McCutchan, P. Gomes, R. Camacho, and N. Taveira.** 2011. Origin and Epidemiological History of HIV-1 CRF14_BG. *PLoS ONE* **6**: e24130.
10. **Barré-Sinoussi, F., J. C. Chermann, F. Rey, M. T. Nugeyre, S. Chamaret, J. Gruest, C. Dautuet, C. Axler-Blin, F. Vézinet-Brun, C. Rouzioux, W. Rozenbaum,**

- and L. Montagnier.** 1983. Isolation of a T-Lymphotropic Retrovirus from a Patient at Risk for Acquired Immune Deficiency Syndrome (AIDS). *Science*. **220**: 868-871.
- 11. Berger, E. A., R. W. Doms, E. M. Fenyö, B. T. M. Korber, D. R. Littman, J. P. Moore, Q. J. Sattentau, H. Schuitemaker, J. Sodroski, and R. A. Weiss.** 1998. A new classification for HIV-1. *Nature*. **391**: 240.
 - 12. Berger, E. A., P. M. Murphy, and J. M. Farber.** 1999. CHEMOKINE RECEPTORS AS HIV-1 CORECEPTORS: Roles in Viral Entry, Tropism, and Disease. *Annu. Rev. Immunol.* **17**: 657-700.
 - 13. Binley, J. M., T. Wrin, B. Korber, M. B. Zwick, M. Wang, C. Chappey, G. Stiegler, R. Kunert, S. Zolla-Pazner, H. Katinger, C. J. Petropoulos, and D. R. Burton.** 2004. Comprehensive Cross-Clade Neutralization Analysis of a Panel of Anti-Human Immunodeficiency Virus Type 1 Monoclonal Antibodies. *J. Virol.* **78**: 13232-13252.
 - 14. Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research*. **7**: 1513-1523.
 - 15. Briz, V., E. Poveda, and V. Soriano.** 2006. HIV entry inhibitors: mechanisms of action and resistance pathways. *J. Antimicrob. Chemother.* **57**: 619-627.
 - 16. Cann, A. J., M. J. Churcher, M. Boyd, W. O'Brien, J. Q. Zhao, J. Zack, and I. S. Y. Chen.** 1992. The Region of the Envelope Gene of Human Immunodeficiency Virus Type 1 Responsible for Determination of Cell Tropism. *J. Virol.* **66**: 305-309.
 - 17. Cantin, R., J. F. Fortin, G. Lamontagne, and M. Tremblay.** 1997. The Acquisition of Host-Derived Major Histocompatibility Complex Class II Glycoproteins by Human Immunodeficiency Virus Type 1 Accelerates the Process of Virus Entry and Infection in Human T-Lymphoid Cells. *Blood*. **90**: 1091-1100.
 - 18. Catasti, P., J. D. Fontenot, E. M. Bradbury, and G. Gupta.** 1994. Local and global structural properties of the HIV-MN V3 Loop. *J. Biol. Chem.* **270**: 2224-2232.
 - 19. CDC.** 1981. Kaposi's sarcoma and Pneumocystis Pneumonia Among Homosexual Men – New York City and California. *MMWR*. **30**: 305-308.
 - 20. CDC.** 1981. Pneumocystis Pneumonia – Los Angeles. *MMWR*. **30**: 250-252.
 - 21. CDC.** 1982. Possible transfusion-associated acquired immune deficiency syndrome (AIDS) – California. *MMWR*. **31**: 652-654.

22. **CDC.** 1982. Epidemiologic Notes and Reports Persistent, Generalized Lymphadenopathy among Homosexual Males. *MMWR.* **31:** 249-251.
23. **CDC.** 1982. Opportunistic Infections and Kaposi's Sarcoma, among Haitians in the United States. *MMWR.* **31:** 353-354.
24. **CDC.** 1982. Unexplained Immunodeficiency and Opportunistic Infections in Infants – New York, New Jersey, California. **31:** 665-667.
25. **Chen, Z., P. Telfer, A. Gettie, P. Reed, L. Zhang, D. D. Ho, and P. A. Marx.** 1996. Genetic Characterization of New West African Simian Immunodeficiency Virus SIVsm: Geographic Clustering of Household-Derived SIV Strains with Human Immunodeficiency Virus Type 2 Subtypes and Genetically Diverse Viruses from a Single Feral Sooty Mangabey Troop. *J. Virol.* **70:** 3617-3627.
26. **Chen, Z., A. Luckay, D. L. Sodora, P. Telfer, P. Reed, A. Gettie, J. M. Kanu, R. F. Sadek, J. Yee, D. D. Ho, L. Zhang, and P. A. Marx.** 1997. Human Immunodeficiency Virus Type 2 (HIV-2) Seroprevalence and Characterization of a Distinct HIV-2 Genetic Subtype from the Natural Range of Simian Immunodeficiency Virus-Infected Sooty Mangabeys. *J. Virol.* **71:** 3953-3960.
27. **Clapham, P. R., and A. McKnight.** 2001. HIV-1 receptors and cell tropism. *Br. Med. Bull.* **58:** 43-59.
28. **Clapham, P. R., and A. McKnight.** 2002. Cell surface receptors, virus entry and tropism of primate lentiviruses. *J. Gen. Virol.* **83:** 1809-1829.
29. **Cohen, E. A., E. F. Terwilliger, J. G. Sodroski, and W. A. Haseltine.** 1988. Identification of a protein encoded by the vpu gene of HIV-1. *Nature.* **334:** 532-534.
30. **Cohen, C., E. DeJesus, A. Mills, G. Pierone, Jr., P. Kumar, P. Ruane, R. Elion, G. Fusco, R. Levy, K. Solomon, C. B. Xue, and S. Erickson-Viitanen.** 2007. Potent antiretroviral activity of the once-daily CCR5 antagonist INCB009471 over 14 days of monotherapy. Abstr. TUAB106. *In* 4th IAS Conference on HIV Pathogenesis, Treatment and Prevention, Sydney, Australia.
31. **Connor, R. I., K. E. Sheridan, D. Ceradini, S. Choe, and N. R. Landau.** 1997. Change in Coreceptor Use Correlates with Disease Progression in HIV-1–Infected Individuals. *J. Exp. Med.* **185:** 621-628.
32. **Corbeau, P., and J. Reynes.** 2009. CCR5 antagonism in HIV infection: ways, effects, and side effects. *AIDS.* **23:** 1931-1943.

33. **Corbet, S., M. C. Muller-Trutwin, P. Versmsse, S. Delarue, A. Ayoub, J. Lewis, S. Brunak, P. Martin, F. Brun-Vezinet, F. Simon, F. Barré-Sinoussi, and P. Maucere.** 2000. *env* Sequences of Simian Immunodeficiency Viruses from Chimpanzees in Cameroon Are Strongly Related to Those of Human Immunodeficiency Virus Group N from the Same Geographic Area. *J. Virol.* **74**: 529-534.
34. **Dagleish, A. G., P. C. L. Beverley, P. R. Clapham, D. H. Crawford, M. F. Greaves, and R. A. Weiss.** 1984. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature.* **312**: 763-767.
35. **Damond, F., M. Worobey, P. Campa, I. Farfara, G. Colin, S. Matheron, F. Brun-Vézinet, D. L. Robertson, and F. Simon.** 2004. Identification of a Highly Divergent HIV Type 2 and Proposal for a Change in HIV Type 2 Classification. *AIDS Res. Hum. Retroviruses.* **20**: 666-672.
36. **De Cock, K. M., H. W. Jaffe, and J. W. Curran.** 2011. Reflections on 30 years of AIDS. *Emerg. Infect. Dis.* **17**: 1044-1048.
37. **De Jong, J. J., A. De Ronde, W. Keulen, M. Tersmette, and J. Goudsmit.** 1992. Minimal Requirements for the Human Immunodeficiency Virus Type 1 V3 Domain To Support the Syncytium-Inducing Phenotype: Analysis by Single Amino Acid Substitution. *J. Virol.* **66**: 6777-6780.
38. **De Leys, R., B. Vanderborght, M. V. Haesevelde, L. Heyndrickx, A. van Geel, C. Wauters, R. Bernaerts, E. Saman, P. Nijs, B. Willems, H. Taelman, G. van der Groen, P. Piot, T. Tersmette, J. G. Huisman, and H. Van Heuverswyn.** 1990. Isolation and Partial Characterization of an Unusual Human Immunodeficiency Retrovirus from Two Persons of West-Central African Origin. *J. Virol.* **64**: 1207-1216.
39. **de Oliveira, T., K. Deforche, S. Cassol, M. Salminen, D. Paraskevis, C. Seebregts, J. Snoeck, E. J. van Rensburg, A. M. J. Wensing, D. A. van de Vijver, C. A. Boucher, R. Camacho, and A. M. Vandamme.** 2005. An automated genotyping system for analysis of HIV-1 and other microbial sequences. *Bioinformatics.* **21**: 3797-3800.
40. **Delgado, E., M. M. Thomson, M. L. Villahermosa, M. Sierra, A. Ocampo, C. Miralles, R. Rodríguez-Pérez, J. Diz-Aren, R. Ojea-de Castro, E. Losada, M. T. Cuevas, E. Vázquez-de Parga, R. Carmona, L. Pérez-Álvarez, L. Medrano, L. Cuevas, J. A. Taboada, and R. Nájera.** 2002. Identification of a Newly Characterized

HIV-1 BG Intersubtype Circulating Recombinant Form in Galicia, Spain, Which Exhibits a Pseudotype-Like Virion Structure. *J. Acquir. Immune. Defic. Syndr.* **29**: 536-543.

41. Delwart, E. L., E. G. Shpaer, J. Louwagie, F. E. McCutchan, M. Grez, H. Rubsamen-Waigmann, and J. I. Mullins. 1993. Genetic relationships determined by a DNA heteroduplex mobility assay: analysis of HIV-1 env genes. *Science.* **262**: 1257-1261.

42. Delwart, E. L., B. Herring, A. G. Rodrigo, and J. I. Mullins. 1995. Genetic Subtyping of Human Immunodeficiency Virus Using a Heteroduplex Mobility Assay. *PCR Methods Appl.* **4**: S202-S216.

43. Deng, H., R. Liu, W. Ellmeier, S. Choe, D. Unutmaz, M. Burkhart, P. D. Marzio, S. Marmon, R. E. Sutton, C. M. Hill, C. B. Davis, S. C. Peiper, T. J. Schall, D. R. Littman, and N. R. Landau. 1996. Identification of a major co-receptor for primary isolates of HIV-1. *Nature.* **381**: 661-666.

44. Dorr, P., M. Westby, S. Dobbs, P. Griffin, B. Irvine, M. McCartney, J. Mori, G. Rickett, C. Smith-Burchnell, C. Napier, R. Webster, D. Armour, D. Price, B. Stammen, A. Wood, and M. Perros. 2005. Maraviroc (UK-427,857), a Potent, Orally Bioavailable, and Selective Small-Molecule Inhibitor of Chemokine Receptor CCR5 with Broad-Spectrum Anti-Human Immunodeficiency Virus Type 1 Activity. *Antimicrob. Agents Chemother.* **49**: 4721-4732.

45. Duque, V., Á. Holguín, M. Silvestre, J. González-Lahoz and, V. Soriano. 2003. Human immunodeficiency virus type 1 recombinant B/G subtypes circulating in Coimbra, Portugal. *Clin. Microbiol. Infect.* **9**: 422-425.

46. Eigen, M. 1993. Viral Quasispecies. *Sci. Am.* **269**: 42-49.

47. Esnouf, R., J. Ren, C. Ross, Y. Jones, D. Stammers, and D. Stuart. 1995. Mechanism of inhibition of HIV-1 reverse transcriptase by non-nucleoside inhibitors. *Nat. Struct. Biol.* **2**: 303-308.

48. Esteves, A., R. Parreira, T. Venenno, M. Franco, J. Piedade, J. Germano de Sousa, and W. F. Canas-Ferreira. 2002. Molecular epidemiology of HIV type 1 infection in Portugal: High prevalence of non-B subtypes. *AIDS Res. Hum. Retroviruses.* **18**: 313-325.

49. **Esteves, A., R. Parreira, J. Piedade, T. Venenno, M. Franco, J. Germano de Sousa, L. Patrício, P. Brum, A. Costa, and W. F. Canas-Ferreira.** 2003. Spreading of HIV-1 Subtype G and *envB/gagG* Recombinant Strains among Injecting Drug Users in Lisbon, Portugal. *AIDS Res. Hum. Retroviruses*. **19**: 511-517.
50. **Etemad, B., A. Fellows, B. Kwambana, A. Kamat, Y. Feng, S. Lee, and M. Sagar.** 2009. Human Immunodeficiency Virus Type 1 V1-to-V5 Envelope Variants from the Chronic Phase of Infection Use CCR5 and Fuse More Efficiently than Those from Early after Infection. *J. Virol*. **83**: 9694-9708.
51. **European Medicine Agency.** 2012. Celsentri: European public assessment report –Product Information. [http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/000811/WC500022190.pdf, acesso Junho 2012].
52. **Fais, S., M. R. Capobianchi, I. Abbate, C. Castilletti, M. Gentile, P. Cordiali Fei, F. Ameglio, and F. Dianzani.** 1995. Unidirectional budding of HIV-1 at the site of cell-to-cell contact is associated with co-polarization of intercellular adhesion molecules and HIV-1 viral matrix protein. *AIDS*. **9**: 329-335.
53. **Fanales-Belasio, E., M. Raimondo, B. Suligoi, and S. Buttò.** 2010. HIV virology and pathogenetic mechanisms of infection: a brief overview. *Ann. Ist. Super. Sanità*. **46**: 5-14.
54. **Feinberg, M. B., R. F. Jarret, A. Aldovini, R. C. Gallo, and F. Wong-Staal.** 1986. HTLV-III expression and production involve complex regulation at levels of splicing and translation of viral RNA. *Cell*. **46**: 807- 817.
55. **Feng, Y., C. C. Broder, P. E. Kennedy, and E. A. Berger.** 1996. HIV-1 Entry Cofactor: Functional cDNA Cloning of a Seven-Transmembrane, G Protein-Coupled Receptor. *Science*. **272**: 872-877.
56. **Fortin, J. F., R. Cantin, G. Lamontagne, and M. Tremblay.** 1997. Host-Derived ICAM-1 Glycoproteins Incorporated on Human Immunodeficiency Virus Type 1 Are Biologically Active and Enhance Viral Infectivity. *J. Virol*. **71**: 3588-3596.
57. **Fouchier, R. A. M., M. Groenink, N. A. Kootstra, M. Tersmette, H. G. Huisman, F. Miedema, and H. Schuitemaker.** 1992. Phenotype-Associated Sequence Variation in the Third Variable Domain of the Human Immunodeficiency Virus Type 1 gp120 Molecule. *J. Virol*. **66**: 3183-3187.

58. Frankel, A. D., and J. A. T. Young. 1998. HIV-1: Fifteen Proteins and an RNA. *Annu. Rev. Biochem.* **67**: 1-25.
59. Freed, E. O. 2001. HIV-1 Replication. *Somat. Cell. Mol. Genet.* **26**: 13-33.
60. Gallo, R., F. Wong-Staal, L. Montagnier, W. A. Haseltine, and M. Yoshida. 1988. HIV/HTLV gene nomenclature. *Nature.* **333**: 504.
61. Gao, F., L. Yue, A. T. White, P. G. Pappas, J. Barchue, A. P. Hanson, B. M. Greene, P. M. Sharp, G. M. Shaw, and B. H. Hahn. 1992. Human infection by genetically diverse SIVsm –related HIV-2 in West Africa. *Nature.* **358**: 495-499.
62. Gao, F., L. Yue, D. L. Robertson, S. C. Hill, H. Hui, R. J. Biggar, A. E. Neequaye, T. M. Whelan, D. D. Ho, G. M. Shaw, P. M. Sharp, and B. H. Hahn. 1994. Genetic Diversity of Human Immunodeficiency Virus Type 2: Evidence for Distinct Sequence Subtypes with Differences in Virus Biology. *J. Virol.* **68**: 7433-7447.
63. Gao, F., E. Bailes, D. L. Robertson, Y. Chen, C. M. Rodenburg, S. F. Michael, L. B. Cummins, L. O. Arthur, M. Peeters, G. M. Shaw, P. M. Sharp, and B. H. Hahn. 1999. Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. *Nature.* **397**: 436-441.
64. Gao, F., N. Vidal, Y. Li, S. A. Trask, Y. Chen, L. G. Kostrikis, D. D. Ho, J. Kim, M. D. Oh, K. Choe, M. Salminen, D. L. Robertson, G. M. Shaw, B. H. Hahn, and M. Peeters. 2001. Evidence of Two Distinct Subsubtypes within the HIV-1 Subtype A Radiation. *AIDS Res. Hum. Retroviruses.* **17**: 675-688.
65. Garrido, C., V. Roulet, N. Chueca, E. Poveda, A. Aguilera, K. Skrabal, N. Zahonero, S. Carlos, F. García, J. L. Faudon, V. Soriano, and C. de Mendoza. 2008. Evaluation of Eight Different Bioinformatics Tools To Predict Viral Tropism in different Human Immunodeficiency Virus Type 1 Subtypes. *J. Clin. Microbiol.* **46**: 887-891.
66. Gartner, S., P. Markovits, D. M. Markovitz, M. H. Kaplan, R. C. Gallo, and M. Popovic. 1986. The role of mononuclear phagocytes in HTLV-III/LAV infection. *Science.* **233**: 215-219.
67. Geijtenbeek, T. B. H., D. S. Kwon, R. Torensma, S. J. van Vliet, G. C. F. van Duinhoven, J. Middel, I. L. M. H. A. Cornelissen, H. S. L. M. Nottet, V. N. KewalRamani, D. R. Littman, C. G. Figdor, and Y. van Kooyk. 2000. DC-SIGN, a

Dendritic Cell-Specific HIV-1-Binding Protein that Enhances *trans*-Infection of T Cells. *Cell*. **100**: 587-597.

68. Gelderblom, H., H. Reupke, T. Winkel, R. Kunze, and G. Pauli. 1987. MHC-antigens: constituents of the envelopes of human and simian immunodeficiency viruses. *Z. Naturforsch. C*. **42**: 1328-1334.

69. Gilliam, B. L., D. J. Riedel, and R. R. Redfield. 2010. Clinical use of CCR5 inhibitors in HIV and beyond. *J. Transl. Med.* **9**: S1-S9.

70. Gomez, C., and T. J. Hope. 2005. The ins and outs of HIV replication. *Cell Microbiol.* **7**: 621-626.

71. Gottlieb, M. S., R. Schroff, H. M. Schanker, J. D. Weisman, P. T. Fan, R. A. Wolf, and A. Saxon. 1981. *Pneumocystis carinii* pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency. *N. Engl. J. Med.* **305**: 1425-1431.

72. Gürtler, L. G., P. H. Hauser, J. Eberle, A. von Brunn, S. Knapp, L. Zekeng, J. M. Tsague, and L. Kaptue. 1994. A New Subtype of Human Immunodeficiency Virus Type 1 (MVP-5180) from Cameroon. *J. Virol.* **68**: 1581-1585.

73. Hahn, B. H., G. M. Shaw, M. E. Taylor, R. R. Redfield, P. D. Markham, S. Z. Salahuddin, F. Wong-Staal, R. C. Gallo, E. S. Parks, and W. P. Parks. 1986. Genetic variation in HTLV-III/LAV over time in patients with AIDS or at risk for AIDS. *Science*. **232**: 1548-1553.

74. Hahn, B. H., G. M. Shaw, K. M. De Cock, and P. M. Sharp. 2000. AIDS as a Zoonosis: Scientific and Public Health Implications. *Science*. **287**: 607-614.

75. Hartley, O., H. Gaertner, J. Wilken, D. Thompson, R. Fish, A. Ramos, C. Pastore, B. Dufour, F. Cerini, A. Melotti, N. Heveker, L. Picard, M. Alizon, D. Mosier, S. Kent, and R. Offord. 2004. Medicinal chemistry applied to a synthetic protein: Development of highly potent HIV entry inhibitors. *Proc. Natl. Acad. Sci. USA*. **101**: 16460-16465.

76. Hazuda, D. J., P. Felock, M. Witmer, A. Wolfe, K. Stillmock, J. A. Grobler, A. Espeseth, L. Gabryelski, W. Schleif, C. Blau, and M. D. Miller. 2000. Inhibitors of Strand Transfer That Prevent Integration and Inhibit HIV-1 Replication in Cells. *Science*. **287**: 646-650.

77. **Hoxie, J. A., B. S. Haggarty, S. E. Bonser, J. L. Rackowski, H. Shan, and P. J. Kanki.** 1988. Biological Characterization of a Simian Immunodeficiency Virus-Like Retrovirus (HTLV-IV): Evidence for CD4-Associated Molecules Required for Infection. *J. Virol.* **62**: 2557-2568.
78. **Hungnes, O., T. O. Jonassen, C. M. Jonassen, and B. Grinde.** 2000. Molecular epidemiology of viral infections. How sequence information helps us understand the evolution and dissemination of viruses. *APMIS.* **108**: 81-97.
79. **Hwang, S. S., T. J. Boyle, H. K. Lyerly, and B. R. Cullen.** 1991. Identification of the envelope V3 loop as the primary determinant of cell tropism in HIV-1. *Science.* **253**: 71-74.
80. **Ibe, S., Y. Yokomaku, T. Shiino, R. Tanaka, J. Hattori, S. Fujisaki, Y. Iwatani, N. Mamiya, M. Utsumi, S. Kato, M. Hamaguchi, and W. Sugiura.** 2010. HIV-2 CRF01_AB: First Circulating Recombinant Form of HIV-2. *J. Acquir. Immune. Defic. Syndr.* **54**: 241-247.
81. **INSA.** 2011. Infecção VIH/SIDA EM Portugal. A situação em 31 de Dezembro de 2010. Departamento de Doenças Infecciosas. Unidade de Referencia e Vigilância Epidemiológica. Núcleo de Vigilância Laboratorial de Doenças Infecciosas. Instituto Nacional de Saúde Dr. Ricardo Jorge. [http://www.insa.pt/sites/INSA/Português/Publicacoes/Outros/Documents/DoencasInfecciosas/Documento_VIH_SIDA_142_31_12_2010.pdf, acesso Agosto 2012].
82. **Jacobson, J. M., M. S. Saag, M. A. Thompson, M. A. Fischl, R. Liporace, R. C. Reichman, R. R. Redfield, C. J. Fichtenbaum, B. S. Zingman, M. C. Patel, J. D. Murga, S. M. Pemrick, P. D'Ambrosio, M. Michael, H. Kroger, H. Ly, Y. Rotshteyn, R. Buice, S. A. Morris, J. J. Stavola, P. J. Maddon, A. B. Kremer, and W. C. Olson.** 2008. Antiviral Activity of Single-Dose PRO 140, a CCR5 Monoclonal Antibody, in HIV-Infected Adults. *J. Infect. Dis.* **198**: 1345-1352.
83. **Janssens, W., L. Heyndrickx, K. Fransen, J. Motte, M. Peeters, J. N. Nkengasong, P. M. Ndumbe, E. Delaporte, J. L. Perret, C. Atende, P. Piot, and G. Van Der Groen.** 1994. Genetic and Phylogenetic Analysis of env Subtypes G and H in Central Africa. *AIDS Res. Hum. Retroviruses.* **10**: 877-879.
84. **Jensen, M. A., F. S. Li, A. B. van 't Wout, D. C. Nickle, D. Shriner, H. X. He, S. McLaughlin, R. Shankarappa, J. B. Margolick, and J. I. Mullins.** 2003.

Improved Coreceptor Usage Prediction and Genotypic Monitoring of R5-to-X4 Transition by Motif Analysis of Human Immunodeficiency Virus Type 1 *env* V3 Loop Sequences. *J. Virol.* **77**: 13376-13388.

85. Jensen, M. A., M. Coetzer, A. B. van 't Wout, L. Morris, and J. I. Mullins. 2006. A Reliable Phenotype Predictor for Human Immunodeficiency Virus Type 1 Subtype C Based on Envelope V3 Sequences. *J. Virol.* **80**: 4698-4704.

86. Jern, P., R. A. Russell, V. K. Pathak, and J. M. Coffin. 2009. Likely Role of APOBEC3G-Mediated G-to-A Mutations in HIV-1 Evolution and Drug Resistance. *PLoS Pathog.* **5**: e1000367.

87. Jetzt, A. E., H. Yu, G. J. Klarmann, Y. Ron, B. D. Preston, and J. P. Dougherty. 2000. High Rate of Recombination throughout the Human Immunodeficiency Virus Type 1 Genome. *J. Virol.* **74**: 1234-1240.

88. Kan, N. C., G. Franchini, F. Wong-Staal, G. C. DuBois, W. G. Robey, J. A. Lautenberger, and T. S. Papas. 1986. Identification of HTLV-III/LAV *src* gene product and detection of antibodies in human sera. *Science.* **231**: 1553-1555.

89. Kandathil, A. J., S. Ramalingam, R. Kannangai, S. David, and G. Sridharan. 2005. Molecular epidemiology of HIV. *Indian J. Med. Res.* **121**: 333-344.

90. Kilgore, N. R., K. Salzwedel, M. Reddick, G. P. Allaway, and C. T. Wild. 2003. Direct Evidence that C-Peptide Inhibitors of Human Immunodeficiency Virus Type 1 Entry Bind to the gp41 N-Helical Domain in Receptor-Activated Viral Envelope. *J. Virol.* **77**: 7669-7672.

91. Kimura, M. 1980. A Simple Method for Estimating Evolutionary Rates of Base Substitutions Through Comparative Studies of Nucleotide Sequences. *J. Mol. Evol.* **16**: 111-120.

92. Korber, B., M. Muldoon, J. Theiler, F. Gao, R. Gupta, A. Lapedes, B. H. Hahn, S. Wolinsky, and T. Bhattacharya. 2000. Timing the Ancestor of the HIV-1 Pandemic Strains. *Science.* **288**: 1789-1796.

93. Kwong, P. D., R. Wyatt, J. Robinson, R. W. Sweet, J. Sodroski, and W. A. Hendrickson. 1998. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature.* **393**: 648-659.

- 94. Lal, R. B., S. Chakrabarti, and C. Yang.** 2005. Impact of genetic diversity of HIV-1 on diagnosis, antiretroviral therapy and vaccine development. *Indian J. Med Res.* **121**: 287-314.
- 95. Lalezari, J., G. K. Yadavalli, M. Para, G. Richmond, E. DeJesus, S. J. Brown, W. Cai, C. Chen, J. Zhong, L. A. Novello, M. M. Lederman, and G. M. Subramanian.** 2008. Safety, Pharmacokinetics, and Antiviral Activity of HGS004, a Novel Fully Human IgG4 Monoclonal Antibody against CCR5, in HIV-1–Infected Patients. *J. Infect. Dis.* **197**: 721-727.
- 96. Lalezari, J., J. Gathe, C. Brinson, M. Thompson, C. Cohen, E. Dejesus, J. Galindez, J. A. Ernst, D. E. Martin, and S. M. Paleja.** 2011. Safety, Efficacy, and Pharmacokinetics of TBR-652, a CCR5/CCR2 Antagonist, in HIV-1–Infected, Treatment-Experienced, CCR5 Antagonist–Naive Subjects. *J. Acquir. Immune. Defic. Syndr.* **57**: 118-125.
- 97. Land, A. and I. Braakman.** 2001. Folding of the human immunodeficiency virus type 1 envelope glycoprotein in the endoplasmic reticulum. *Biochimie.* **83**: 783-790.
- 98. Larkin, M., R. A. Childs, T. J. Matthews, S. Thiel, T. Mizuochi, A. M. Lawson, J. S. Savill, C. Haslett, R. Diaz, and T. Feizi.** 1989. Oligosaccharide-mediated interactions of the envelope glycoprotein gp120 of HIV-1 that are independent of CD4 recognition. *AIDS.* **3**: 793-798.
- 99. Lee, W. R., W. J. Syu, B. Du, M. Matsuda, S. Tan, A. Wolf, M. Essex, and T. H. Lee.** 1992. Nonrandom distribution of gp120 N-linked glycosylation sites important for infectivity of human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci. USA.* **89**: 2213-2217.
- 100. Leonard, C. K., M. W. Spellman, L. Riddle, R. J. Harris, J. N. Thomas, and T. J. Gregory.** 1990. Assignment of Intrachain Disulfide Bonds and Characterization of Potential Glycosylation Sites of the Type 1 Recombinant Human Immunodeficiency Virus Envelope Glycoprotein (gp120) Expressed in Chinese Hamster Ovary Cells. *J. Biol. Chem.* **265**: 10373-10382.
- 101. Lewis, M., P. Simpson, C. Delogne, J. Sullivan, H. Valdez, R. Harrigan, and M. Westby.** 2010. Genotypic Analysis of the HIV-1 gp120 V3 Loop for Treatment-Experienced Patients Enrolled Into the MOTIVATE Studies and Who Received

Maraviroc + Optimized Background Therapy. Abstr.539. *In* Abstracts of the Seventeenth Conference on Retroviruses and Opportunistic Infections, San Francisco. Foundation for Retrovirology and Human Health, Alexandria, VA, USA.

102. Lin, P. F., W. Blair, T. Wang, T. Spicer, Q. Guo, N. Zhou, Y. F. Gong, H. G. H. Wang, R. Rose, G. Yamanaka, B. Robinson, C. B. Li, R. Fridell, C. Deminie, G. Demers, Z. Yang, L. Zadjura, N. Meanwell, and R. Colonno. 2003. A small molecule HIV-1 inhibitor that targets the HIV-1 envelope and inhibits CD4 receptor binding. *Proc. Natl. Acad. Sci. USA.* **100**: 11013-11018.

103. Liu, R., W. A. Paxton, S. Choe, D. Ceradini, S. R. Martin, R. Horuk, M. E. MacDonald, H. Stuhlmann, R. A. Koup, and N. R. Landau. 1996. Homozygous Defect in HIV-1 Coreceptor Accounts for Resistance of Some Multiply-Exposed Individuals to HIV-1 Infection. *Cell.* **86**: 367-377.

104. Loftin, L. M., M. Kienzle, Y. Yi, and R. G. Collman. 2010. R5X4 HIV-1 coreceptor use in primary target cells: implications for coreceptor entry blocking strategies. *J. Transl. Med.* **9**: S3.

105. Lole, K. S., R. C. Bollinger, R. S. Paranjape, D. Gadkari, S. S. Kulkarni, N. G. Novak, R. Ingersoll, H. W. Sheppard, and S. C. Ray. 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J. Virol.* **73**: 152-160.

106. Mansky, L. M., and H. M. Temin. 1995. Lower In Vivo Mutation Rate of Human Immunodeficiency Virus Type 1 than That Predicted from the Fidelity of Purified Reverse Transcriptase. *J. Virol.* **69**: 5087-5094.

107. Marier, J. F., M. Trinh, L. H. Pheng, S. M. Paleja, and D. E. Martin. 2011. Pharmacokinetics and Pharmacodynamics of TBR-652, a Novel CCR5 Antagonist, in HIV-1-Infected, Antiretroviral Treatment-Experienced, CCR5 Antagonist-Naïve Patients. *Antimicrob. Agents Chemother.* **55**: 2768-2774.

108. Masur, H., M. A. Michelis, J. B. Greene, I. Onorato, R. A. Stouwe, R. S. Holzman, G. Wormser, L. Brettman, M. Lange, H. W. Murray, and S. Cunningham-Rundles. 1981. An Outbreak of Community-Acquired *Pneumocystis carinii* Pneumonia - Initial Manifestation of Cellular Immune Dysfunction. *N. Engl. J. Med.* **305**: 1431-1438.

109. **McGovern, R. A., A. Thielen, T. Mo, W. Dong, C. K. Woods, D. Chapman, M. Lewis, I. James, J. Heera, H. Valdez, and P. R. Harrigan.** 2010. Population-based V3 genotypic tropism assay: a retrospective analysis using screening samples from the A4001029 and MOTIVATE studies. *AIDS*. **24**: 2517-2525.
110. **Meloni, S. T., B. Kim, J. L. Sankalé, D. J. Hamel, S. Tovanabutra, S. Mboup, F. E. McCutchan, and P. J. Kanki.** 2004. Distinct Human Immunodeficiency Virus Type 1 Subtype A Virus Circulating in West Africa: Sub-Subtype A3. *J. Virol.* **78**: 12438-12445.
111. **Merluzzi, V. J., K. D. Hargrave, M. Labadia, K. Grozinger, M. Skoog, J. C. Wu, C. K. Shih, K. Eckner, S. Hattox, and J. Adams.** 1990. Inhibition of HIV-1 replication by a nonnucleoside reverse transcriptase inhibitor. *Science*. **250**: 1411-1413.
112. **Mitsuya, H., and S. Broder.** 1986. Inhibition of the in vitro infectivity and cytopathic effect of human T-lymphotrophic virus type III/lymphadenopathy-associated virus (HTLV-III/LAV) by 2',3'-dideoxynucleosides. *Proc. Natl. Acad. Sci. USA*. **83**: 1911-1915.
113. **Modrow, S., B. H. Hahn, G. M. Shaw, R. C. Gallo, F. Wong-Staal, and H. Wolf.** 1987. Computer-Assisted Analysis of Envelope Protein Sequences of Seven Human Immunodeficiency Virus Isolates: Prediction of Antigenic Epitopes in Conserved and Variable Regions. *J. Virol.* **61**: 570-578.
114. **Moore, J. P., and J. Sodroski.** 1996. Antibody Cross-Competition Analysis of the Human Immunodeficiency Virus Type 1 gp120 Exterior Envelope Glycoprotein. *J. Virol.* **70**: 1863-1872.
115. **Mori, J., M. Mosley, M. Lewis, P. Simpson, J. Toma, W. Huang, J. Whitcomb, G. Ciaramella, and M. Westby.** 2007. Characterization of Maraviroc Resistance in Patients Failing Treatment with CCR5-Tropic HIV-1 in MOTIVATE 1 and 2. Abstr. 10. *In: XVI International HIV Drug Resistance Workshop*. Barbados, West Indies.
116. **Murakami, T., S. Kumakura, T. Yamazaki, R. Tanaka, M. Hamatake, K. Okuma, W. Huang, J. Toma, J. Komano, M. Yanaka, Y. Tanaka, and N. Yamamoto.** 2009. The Novel CXCR4 Antagonist KRH-3955 Is an Orally Bioavailable and Extremely Potent Inhibitor of Human Immunodeficiency Virus Type 1 Infection: Comparative Studies with AMD3100. *Antimicrob. Agents Chemother.* **53**: 2940-2948.

117. **Nichols, W. G., H. M. Steel, T. Bonny, K. Adkison, L. Curtis, J. Millard, K. Kabeya, and N. Clumeck.** 2008. Hepatotoxicity Observed in Clinical Trials of Atravirine (GW873140). *Antimicrob. Agents Chemother.* **52**: 858-865.
118. **Page, R. D. M., and E. C. Holmes.** 1998. *Molecular Evolution: A Phylogenetic Approach.* Blackwell Science Ltd. Oxford.
119. **Palma, A. C., F. Araújo, V. Duque, F. Borges, M. T. Paixão, and R. Camacho, on behalf of the Portuguese SPREAD Network.** 2007. Molecular epidemiology and prevalence of drug resistance-associated mutations in newly diagnosed HIV-1 patients in Portugal. *Infect. Genet. Evol.* **7**: 391-398.
120. **Perelson A. S., A. U. Neumann, M. Markowitz, J. M. Leonard, and D. D. Ho.** 1996. HIV-1 Dynamics in Vivo: Virion Clearance Rate, Infected Cell Life-Span, and Viral Generation Time. *Science.* **271**: 1582-1586.
121. **Plantier, J. C., M. Leoz, J. E. Dickerson, F. De Oliveira, F. Cordonnier, V. Lemée, F. Damond, D. L. Robertson, and F. Simon.** 2009. A new immunodeficiency virus derived from gorillas. *Nat. Med.* **15**: 871-872.
122. **Poveda, E., V. Briz, M. Quiñones-Mateu, and V. Soriano.** 2006. HIV tropism: diagnostic tools and implications for disease progression and treatment with entry inhibitors. *AIDS.* **20**: 1359-1367.
123. **Poveda, E., V. Briz, V. Roulet, M. del Mar González, J. L. Faudon, K. Skrabal, and V. Soriano.** 2007. Correlation between a phenotypic assay and three bioinformatic tools for determining HIV co-receptor use. *AIDS.* **21**: 1487-1490.
124. **Poveda, E., E. Seclén, M. del M. González, F. García, N. Chueca, A. Aguilera, J. J. Rodríguez, J. González-Lahoz, and V. Soriano.** 2009. Design and validation of new genotypic tools for easy and reliable estimation of HIV tropism before using CCR5 antagonists. *J. Antimicrob. Chemother.* **63**: 1006-1010.
125. **Ragni M. V., J. A. Spero, J. H. Lewis, and F. A. Bontempo.** 1983. Acquired-immunodeficiency- like syndrome in two haemophiliacs. *Lancet.* **1**: 213-214.
126. **Rambaut, A., D. Posada, K. A. Crandall, and E. C. Holmes.** 2004. The Causes and consequences of HIV evolution. *Nat. Rev. Genet.* **5**: 52-61.
127. **Ratner, L. , W. Haseltine, R. Patarca, K. J. Livak, B. Starcich, S. F. Josephs, E. R. Doran, J. A. Rafalski, E. A. Whitehorn, K. Baumeister, L. Ivanoff, S. R. Petteway, JR, M. L. Pearson, J. A. Lautenberger, T. S. Papas, J. Ghayeb, N. T.**

- Chang, R. C. Gallo, and F. Wong-Staal.** 1985. Complete nucleotide sequence of the AIDS virus, HTLV-III. *Nature*. **313**: 277-284.
- 128. Ray, N.** 2008. Maraviroc in the treatment of HIV infection. *Drug Des. Devel. Ther.* **2**: 151-161.
- 129. Raymond, S., P. Delobel, M. Mavigner, M. Cazabat, C. Souyris, K. Sandres-Sauné, L. Cuzin, B. Marchou, P. Massip, and J. Izopet.** 2008. Correlation between genotypic predictions based on V3 sequences and phenotypic determination of HIV-1 tropism. *AIDS*. **22**: F11-F16.
- 130. Recordon-Pinson, P., C. Soulié, P. Flandre, D. Descamps, M. Lazrek, C. Charpentier, B. Montes, M. A. Trabaud, J. Cottalorda, V. Schneider, L. Morand-Joubert, C. Tamalet, D. Desbois, M. Macé, V. Ferré, A. Vabret, A. Ruffault, C. Pallier, S. Raymond, J. Izopet, J. Reynes, A. G. Marcelin, B. Masquelier and the ANRS AC11 Resistance Study Group.** 2010. Evaluation of the genotypic prediction of HIV-1 coreceptor use versus a phenotypic assay and correlation with the virological response to maraviroc: the ANRS GenoTropism Study. *Antimicrob. Agents Chemother.* **54**: 3335-3340.
- 131. Roberts, J. D., K. Bebenek, and T. A. Kunkel.** 1988. The accuracy of reverse transcriptase from HIV-1. *Science*. **242**: 1171-1173.
- 132. Robertson, D. L., J. P. Anderson, J. A. Bradac, J. K. Carr, B. Foley, R. K. Funkhouser, F. Gao, B. H. Hahn, M. L. Kalish, C. Kuiken, G. H. Learn, T. Leitner, F. McClutchan, S. Osmanov, M. Peeters, D. Pieniazek, M. Salminen, P. M. Sharp, S. Wolinsky, and B. Korber.** 2000. HIV-1 nomenclature proposal. *Science*. **288**: 55-56.
- 133. Saag, M. S., B. H. Hahn, J. Gibbons, Y. Li, E. S. Parks, W. P. Parks, and G. M. Shaw.** 1988. Extensive variation of human immunodeficiency virus type-1 *in vivo*. *Nature*. **334**: 440-444.
- 134. Sadler, H. A., M. D. Stenglein, R. S. Harris, and L. M. Mansky.** 2010. APOBEC3G Contributes to HIV-1 Variation through Sublethal Mutagenesis. *J. Virol.* **84**: 7396-7404.
- 135. Saitou, N., and M. Nei.** 1987. The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees. *Mol. Biol. Evol.* **4**: 406-425.

136. **Samson, M., F. Libert, B. J. Doranz, J. Rucker, C. Liesnard, C. M. Farber, S. Saragosti, C. Lapoum  roulie, J. Cognaux, C. Forceille, G. Muyldermans, C. Verhofstede, G. Burtonboy, M. Georges, T. Imai, S. Rana, Y. Yi, R. J. Smyth, R. G. Collman, R. W. Doms, G. Vassart, and M. Parmentier.** 1996. Resistance to HIV-1 infection in Caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature*. **382**: 722-725.
137. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA*. **74**: 5463-5467.
138. **Saracino, A., L. Monno, G. Brindicci, G. Trillo, M. Altamura, G. Punzi, A. Lagioia, and G. Angarano.** 2010. Are the proposed env mutations actually associated with resistance to maraviroc? *J. Acquir. Immune. Defic. Syndr.* **53**: 550-552.
139. **Secl  n, E., M. del M. Gonz  lez, M. Lapaz, C. Rodr  guez, J. del Romero, A. Aguilera, C. de Mendoza, V. Soriano, and E. Poveda.** 2010. Primary resistance to maraviroc in a large set of R5-V3 viral sequences from HIV-1-infected patients. *J. Antimicrob. Chemother.* **65**: 2502-2504.
140. **Sethi, A. K., D. D. Celentano, S. J. Gange, R. D. Moore, and J. E. Gallant.** 2003. Association between Adherence to Antiretroviral Therapy and Human Immunodeficiency Virus Drug Resistance. *Clin. Infect. Dis.* **37**: 1112-1118.
141. **Shioda, T., J. A. Levy, and C. Cheng-Mayer.** 1992. Small amino acid changes in the V3 hypervariable region of gp120 can affect the T-cell-line and macrophage tropism of human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci. USA*. **89**: 9434-9438.
142. **Sierra, S., B. Kupfer, and R. Kaiser.** 2005. Basics of the virology of HIV-1 and its replication. *J. Clin. Virol.* **34**: 233-244.
143. **Sierra, S., R. Kaiser, A. Thielen, and T. Lengauer.** 2007. Genotypic coreceptor analysis. *Eur. J. Med. Res.* **12**: 453-462.
144. **Simon, F., P. Mauc   re, P. Roques, I. Loussert-Ajaka, M. C. M  ller-Trutwin, S. Saragosti, M. C. Georges-Courbot, F. Barr  -Sinoussi, and F. Brun-V  zinet.** 1998. Identification of a new human immunodeficiency virus type 1 distinct from group M and group O. *Nat. Med.* **4**: 1032-1037.

145. **Sodroski, J., W. C. Goh, C. Rosen, A. Dayton, E. Terwilliger, and W. Haseltine.** 1986. A second post-transcriptional *trans*-activator gene required for HTLV-III replication. *Nature*. **321**: 412-417.
146. **Soulié, C., I. Malet, S. Lambert-Niclot, R. Tubiana, M. Thévenin, A. Simon, R. Murphy, C. Katlama, V. Calvez, and A. G. Marcelin.** 2008. Primary genotypic resistance of HIV-1 to CCR5 antagonists in CCR5 antagonist treatment-naïve patients. *AIDS*. **22**: 2212-2214.
147. **Sozzani, S., W. Luini, A. Borsatti, N. Polentarutti, D. Zhou, L. Piemonti, G. D'Amico, C. A. Power, T. N. Wells, M. Gobbi, P. Allavena, and A. Mantovani.** 1997. Receptor expression and responsiveness of human dendritic cells to a defined set of CC and CXC chemokines. *J. Immunol.* **159**: 1993-2000.
148. **Stone, V. E., J. Jordan, J. Tolson, R. Miller, and T. Pilon.** 2004. Perspectives on Adherence and Simplicity for HIV-Infected Patients on Antiretroviral Therapy: Self-Report of the Relative Importance of Multiple Attributes of Highly Active Antiretroviral Therapy (HAART) Regimens in Predicting Adherence. *J. Acquir. Immune. Defic. Syndr.* **36**: 808-816.
149. **Strizki, J. M., C. Tremblay, S. Xu, L. Wojcik, N. Wagner, W. Gonsiorek, R. W. Hipkin, C. C. Chou, C. Pugliese-Sivo, Y. Xiao, J. R. Tagat, K. Cox, T. Priestley, S. Sorota, W. Huang, M. Hirsch, G. R. Reyes, and B. M. Baroudy.** 2005. Discovery and Characterization of Vicriviroc (SCH 417690), a CCR5 Antagonist with Potent Activity against Human Immunodeficiency Virus Type 1. *Antimicrob. Agents Chemother.* **49**: 4911-4919.
150. **Svicher, V., C. Alteri, A. Artese, J. M. Zhang, G. Costa, F. Mercurio, R. D'Arrigo, S. Alcaro, G. Palù, M. Clementi, M. Zazzi, M. Andreoni, A. Antinori, A. Lazzarin, F. Ceccherini-Silberstein, and C. F. Perno on behalf of the OSCAR study group.** 2011. Identification and structural characterization of novel genetic elements in the HIV-1 V3 loop regulating coreceptor usage. *Antivir. Ther.* **16**: 1035-1045.
151. **Temesgen, Z., D. Warnke, and M. J. Kasten.** 2006. Current status of antiretroviral therapy. *Expert. Opin. Pharmacother.* **7**: 1541-1554.
152. **Tersmette, M., R. E. Y. de Goede, B. J. M. Al, I. N. Winkel, R. A. Gruters, H. T. Cuypers, H. G. Huisman, and F. Miedema.** 1988. Differential Syncytium-Inducing Capacity of Human Immunodeficiency Virus Isolates: Frequent Detection of

Syncytium-Inducing Isolates in Patients with Acquired Immunodeficiency Syndrome (AIDS) and AIDS-Related Complex. *J. Virol.* **62**: 2026-2032.

153. Tersmette, M., J. J. M. van Dongen, P. R. Clapham, R. E. Y. De Goede, I. L. M. Wolvers-Tettero, A. G. Van Kessel, J. G. Huisman, R. A. Weiss, and F. Miedema. 1989. Human immunodeficiency virus infection studied in CD4-expressing human-murine T-cell hybrids. *Virology.* **168**: 267-273.

154. Thompson, M. M., L. Pérez-Álvarez, and R. Nájera. 2002. Molecular epidemiology of HIV-1 genetic forms and its significance for vaccine development and therapy. *Lancet Infect. Dis.* **2**: 461-471.

155. Tilton, J. C., and R. W. Doms. 2010. Entry inhibitors in the treatment of HIV-1 infection. *Antiviral. Res.* **85**: 91-100.

156. Triques, K., A. Bourgeois, S. Saragosti, N. Vidal, E. Mpoudi-Ngole, N. Nzilambi, C. Apetrei, M. Ekwilanga, E. Delaporte, and M. Peeters. 1999. High Diversity of HIV-1 Subtype F Strains in Central Africa. *Virology.* **259**: 99-109.

157. Triques, K., A. Bourgeois, N. Vidal, E. Mpoudi-Ngole, C. Mulanga-Kabeya, N. Nzilambi, N. Torimiro, E. Saman, E. Delaporte, and M. Peeters. 2000. Near-Full-Length Genome Sequencing of Divergent African HIV Type 1 Subtype F Viruses Leads to the Identification of a New HIV Type 1 Subtype Designated K. *AIDS Res. Hum. Retroviruses.* **16**: 139-151.

158. UNAIDS/WHO. 2010. Global Report. UNAIDS Report on the Global AIDS Epidemic, 2010. [http://www.unaids.org/en/media/unaids/contentassets/documents/unaidspublication/2010/20101123_globalreport_en%5B1%5D.pdf, acesso Maio 2012].

159. Vallari, A., V. Holzmayer, B. Harris, J. Yamaguchi, C. Ngansop, F. Makamche, D. Mbanya, L. Kaptué, N. Ndembi, L. Gürtler, S. Devare, and C. A. Brennan. 2011. Confirmation of Putative HIV-1 Group P in Cameroon. *J. Virol.* **85**: 1403-1407.

160. Van Heuverswyn, F., Y. Li, C. Neel, E. Bailes, B. F. Keele, W. Liu, S. Loul, C. Butel, F. Liegeois, Y. Bienvenue, E. M. Ngolle, P. M. Sharp, G. M. Shaw, E. Delaporte, B. H. Hahn, and M. Peeters. 2006. Human immunodeficiency viruses: SIV infection in wild gorillas. *Nature.* **444**: 164.

161. Vandekerckhove, L. P. R., A. M. J. Wensing, R. Kaiser, F. Brun-Vézinet, B. Clotet, A. De Luca, S. Dressler, F. Garcia, A. M. Geretti, T. Klimkait, K. Korn, B.

- Masquelier, C. F. Perno, J. M. Schapiro, V. Soriano, A. Sönnernborg, A. M. Vandamme, C. Verhofstede, H. Walter, M. Zazzi, and C. A. B. Boucher, on behalf of the European Consensus Group on clinical management of tropism testing.** 2011. European guidelines on the clinical management of HIV-1 tropism testing. *Lancet Infect. Dis.* **11**: 394-407.
- 162. Vidal, N., C. Mulanga, S. E. Bazepeo, F. Lepira, E. Delaporte, and M. Peeters.** 2006. Identification and Molecular Characterization of Subsubtype A4 in Central Africa. *AIDS Res. Hum. Retroviruses.* **22**: 182-187.
- 163. Vidal, N., S. E. Bazepeo, C. Mulanga, E. Delaporte, and M. Peeters.** 2009. Genetic Characterization of Eight Full-Length HIV Type 1 Genomes from the Democratic Republic of Congo (DRC) Reveal a New Subsubtype, A5, in the A Radiation That Predominates in the Recombinant Structure of CRF26_A5U. *AIDS Res. Hum. Retroviruses.* **25**: 823-832.
- 164. Wei, X., J. M. Decker, S. Wang, H. Hui, J. C. Kappes, X. Wu, J. F. Salazar-Gonzalez, M. G. Salazar, J. M. Kilby, M. S. Saag, N. L. Komarova, M. A. Nowak, B. H. Hahn, P. D. Kwong, and G. M. Shaw.** 2003. Antibody neutralization and escape by HIV-1. *Nature.* **422**: 307-312.
- 165. Weiss, R. A.** 2002. HIV Receptors and Cellular Tropism. *IUBMB Life.* **53**: 201-205.
- 166. Weiss, R. A.** 2008. Twenty-five years of human immunodeficiency virus research: successes and challenges. *Clin. Exp. Immunol.* **152**: 201-210.
- 167. Westby, M., M. Lewis, J. Whitcomb, M. Youle, A. L. Pozniak, I. T. James, T. M. Jenkins, M. Perros, and E. van der Ryst.** 2006. Emergence of CXCR4-Using Human Immunodeficiency Virus Type 1 (HIV-1) Variants in a Minority of HIV-1-Infected Patients following Treatment with the CCR5 Antagonist Maraviroc Is from a Pretreatment CXCR4-Using Virus Reservoir. *J. Virol.* **80**: 4909-4920.
- 168. Westby, M., C. Smith-Burchnell, J. Mori, M. Lewis, M. Mosley, M. Stockdale, P. Dorr, G. Ciaramella, and M. Perros.** 2007. Reduced maximal inhibition in phenotypic susceptibility assays indicates that viral strains resistant to the CCR5 antagonist maraviroc utilize inhibitor-bound receptor for entry. *J. Virol.* **81**: 2359-2371.

169. **Whitcomb, J. M., W. Huang, S. Fransen, K. Limoli, J. Toma, T. Wrin, C. Chappey, L. D. B. Kiss, E. E. Paxinos, and C. J. Petropoulos.** 2007. Development and Characterization of a Novel Single-Cycle Recombinant-Virus Assay To Determine Human Immunodeficiency Virus Type 1 Coreceptor Tropism. *Antimicrob. Agents Chemother.* **51**: 566-575.
170. **Willey, R. L., R. A. Rutledge, S. Dias, T. Folks, T. Theodore, C. E. Buckler, and M. A. Martin.** 1986. Identification of conserved and divergent domains within the envelope gene of the acquired immunodeficiency syndrome retrovirus. *Proc. Natl. Acad. Sci. USA.* **83**: 5038-5042.
171. **Wolfs, T. F. W., J. J. De Jong, H. Van den Berg, J. M. G. H. Tijnagel, W. J. A. Krone, and J. Goudsmit.** 1990. Evolution of sequences encoding the principal neutralization epitope of human immunodeficiency virus 1 is host dependent, rapid, and continuous. *Proc. Natl. Acad. Sci. USA.* **87**: 9938-9942.
172. **Wong-Staal, F., P. K. Chanda, and J. Ghrayeb.** 1987. Human Immunodeficiency Virus: the eighth gene. *AIDS Res. Hum. Retroviruses.* **3**: 33-39.
173. **Wyatt, R., and J. Sodroski.** 1998. The HIV-1 Envelope Glycoproteins: Fusogens, Antigens, and Immunogens. *Science.* **280**: 1884-1888.
174. **Yamaguchi, J., S. G. Devare, and C. A. Brennan.** 2000. Identification of a New HIV-2 Subtype Based on Phylogenetic Analysis of Full-Length Genomic Sequence. *AIDS Res. Hum. Retroviruses.* **16**: 925-930.
175. **Yerly, S., L. Kaiser, E. Race, J. P. Bru, F. Clavel, and L. Perrin.** 1999. Transmission of antiretroviral-drug-resistant HIV-1 variants. *Lancet.* **354**: 729-733.
176. **Yu, H., A. E. Jetzt, Y. Ron, B. D. Preston, and J. P. Dougherty.** 1998. The Nature of Human Immunodeficiency Virus Type 1 Strand Transfers. *J. Biol. Chem.* **273**: 28384-28391.
177. **Zhang, M., B. Gaschen, W. Blay, B. Foley, N. Haigwood, C. Kuiken and B. Korber.** 2004. Tracking global patterns of *N*-linked glycosylation site variation in highly variable viral glycoproteins: HIV, SIV, and HCV envelopes and influenza hemagglutinin. *Glycobiology.* **14**: 1229-1246.

Anexo 1: Elemento Bibliográfico

MICRO BIOTEC RESEARCH & DEVELOPMENT

³ Laboratório de Referência da Sida, Instituto Nacional de Saúde Dr. Ricardo Jorge, Lisboa

Material and Methods

The flowchart illustrates the process of HIV-1 RNA extraction, RT-PCR, and sequence analysis. It begins with HIV-1 RNA extraction from 52 IDU plasma samples (65% male, mean age: 37.3 years), represented by a yellow liquid in a tube. This is followed by RT-PCR amplification using the Titan One Tube RT-PCR System (Roche, Germany) followed by a second amplification using *Illustra™ puffer Taq Ready-To-Go PCR Beads* (GE Healthcare, England), with Heteroduplex Mobility Assay *env* primers and conditions, shown as a PCR machine and a tube. The next step is DNA sequencing from PCR amplicons or cloned products on pGEM-T Easy vector (Promega, USA), depicted by a DNA sequence logo and a plasmid. The workflow then branches into two main analysis paths. The left path involves Phylogenetic analysis for subtype determination, performed with MEGA 4 (Neighbor-Joining, K2P model, bootstrap of 1,000 replicates), leading to a phylogenetic tree. The right path involves Sequence alignment with ClustalW2¹, after chromatogram analysis and sequence edition with BioEdit v7.0.9.0², leading to a sequence alignment view. Both paths converge on Protein sequence alignment to a consensus B sequence (Los Alamos Database) with ClustalW2¹, after translation by TransSeq³, which leads to a protein sequence alignment view. Finally, the workflow concludes with Coreceptor prediction using PSSM⁴ and gen2phen⁵ programs, and 11/25¹⁸ and V3 charge¹¹ rules. Analyses of polymorphisms are noted as of MVC resistance-associated mutations.

```

graph LR
    A[HIV-1 RNA extraction from 52 IDU plasma samples (65% male, mean age: 37.3 years)] --> B[RT-PCR amplification with Titan One Tube RT-PCR System (Roche, Germany) followed by a second amplification using Illustra™ puffer Taq Ready-To-Go PCR Beads (GE Healthcare, England), with Heteroduplex Mobility Assay env primers and conditions]
    B --> C[DNA sequencing from PCR amplicons or cloned products on pGEM-T Easy vector (Promega, USA)]
    C --> D[Phylogenetic analysis, for subtype determination, performed with MEGA 4 (Neighbor-Joining, K2P model, bootstrap of 1,000 replicates)]
    C --> E[Sequence alignment with ClustalW21, after chromatogram analysis and sequence edition with BioEdit v7.0.9.02]
    D --> F[Protein sequence alignment to a consensus B sequence (Los Alamos Database) with ClustalW21, after translation by TransSeq3]
    E --> F
    F --> G[Coreceptor prediction using PSSM4 and gen2phen5 programs, and 11/2518 and V3 charge11 rules. Analyses of polymorphisms are of MVC resistance-associated mutations.]
  
```

Phylogenetic tree showing relationships between various taxa. The tree is rooted on the left and branches to the right. Taxa are labeled with names like PT_T184A ENV, PT_TX16 ENV, etc. Bootstrap values are shown at the nodes. A pie chart on the right shows the distribution of taxa into four categories: B (blue), A (red), C (green), and D (yellow). The pie chart is divided into segments labeled 1, 2, 3, and 4.

Figure 2. Graphic representation of the phylogenetic tree.

Color	Number of People
Purple	11
Green	11
Red	3
Blue	1

Diagram illustrating the structure of the G15_P161NSA protein, showing various mutations indicated by arrows:

- 16A
- 18G
- G18DEL
- 19S/T
- 20L/F
- 21I
- 22T
- N24_I25INSI
- 25D
- 26V
- 2I
- 4L
- 11R/S
- 13H/S

A 3D pie chart illustrating the distribution of responses for the question 'How often do you use the Internet?'. The chart is divided into four segments labeled 1, 2, 3, and 6. Segment 1 (purple) is the largest, followed by segment 6 (orange), segment 3 (green), and segment 2 (white).

Category	Percentage
1	55.56%
2	7.78%
3	16.67%
6	19.99%

Color	Number of People
Grey	3
Green	2
Brown	3
Pink	28
Red	16
Purple	2
Yellow	1
Orange	1
Cyan	6
Blue	11
Magenta	13

Conclusions

Conclusions

- ♦ It was observed a predominance of HIV-1 subtype B strains (60%) in IDUs from the Greater Lisbon; however, there were also subtypes A, G, F1 and CRF14_BG (Figs. 1 and 2). This finding supports the formerly reported high diversity of HIV-1 subtypes in Portugal¹².
- ♦ The genotypic methods for coreceptor prediction still present limitations, especially when analyzing non-B strains (see Fig. 3, with discordant results). Anyway, these methods are viable alternatives when phenotypic testing is not possible.
- ♦ Several single mutations associated with MVC resistance were observed in our samples, as well as two mutational patterns (Figs. 4 and 5). The presence of these mutations in MVC-naïve patients is a result of the high variability of the V3 region in the gp120 of HIV-1.
- ♦ The impact of the presence of these mutations at baseline of MVC treatment is still unknown, but may have important clinical implications.

References

1. Westby, M. *et al.* (2006), *J. Virol.*, 80: #909-20.
2. Westby, M. & E. Byrd (2010), *Antivir. Chem. Chemother.*, 20: 179-92.
3. Delmar, E.L. *et al.* (1995), *PCR Methods Appl.*, 4: S202-16.
4. <http://www.ncbi.nlm.nih.gov/Genbank/FASTA.html>
5. Hall, T.A. (1999), *Nucl. Acids. Symp. Ser.*, 41: 95-8.
6. Tamura, K. *et al.* (2007), *Mol. Biol. Evol.*, 24: 1596-9.
7. <http://www.abi.com.cn/Tools/emboss/seq/seqmap/>
8. <http://www.ncbi.nlm.nih.gov/Genbank/FASTA.html>
9. <http://careceptor.binf.fr/mpf-imp.mpg.de/index.php>
10. De Jong, J. *et al.* (1992), *J. Virol.*, 76: 3256-60.
11. Feuchter, R. *et al.* (1992), *J. Virol.*, 66: 3183-7.
12. Edwards, A. *et al.* (2002), *AIDS Res. Hum. Immun.*, 15: 109-18.

Anexo 2: Compilação das informações epidemiológicas e clínicas referentes às amostras em estudo

ID	Idade (Anos)	Sexo	Ano de Infecção	País de Origem	CV (cópias de RNA/ml)	Origem dos Parceiros sexuais	Relações Sexuais Não Protegidas	Partilha de seringas	Estádio de Infecção	TAR
TP02	42	F	NR	Portugal	NR	Portuguesa	N	S	A	N
TP03	39	M	NR	Portugal	NR	Portuguesa	S	S	A	S
TP04	42	M	NR	Portugal	NR	NR	S	S	A	S
TP08	37	M	NR	Portugal	NR	NR	S	N	A	S
TP13	40	M	NR	Portugal	NR	Portuguesa	N	N	A	N
TP16	39	M	NR	Portugal	NR	Portuguesa	S	S	A	N
TP19	41	F	NR	Portugal	NR	Portuguesa	N	N	A	S
TP24	38	M	NR	Portugal	NR	Várias	N	N	A	S
TP38	36	M	NR	Portugal	NR	Portuguesa	N	N	A	N
TP42	44	M	NR	Portugal	NR	Várias	S	S	NR	S
TP44	28	F	NR	Portugal	NR	Portuguesa	N	N	NR	NR
TP47	33	F	NR	Portugal	NR	Portuguesa	S	N	NR	S
TP55	37	F	NR	Portugal	NR	Portuguesa	S	S	NR	S
TP63	42	M	NR	Angola	NR	Várias	S	N	NR	S
TP71	46	F	NR	Portugal	NR	Portuguesa	S	N	NR	NR
TP72	44	F	NR	Portugal	NR	Portuguesa	S	N	NR	N
TP85	32	M	NR	Angola	NR	Várias	S	S	NR	N
TP98	47	M	NR	Portugal	NR	Várias	S	S	NR	N
TP107	31	F	NR	Portugal	NR	Portuguesa	S	N	NR	N
TP114	38	M	NR	Portugal	NR	Portuguesa	S	N	NR	N
TP128	53	F	NR	Angola	NR	Várias	S	N	NR	S
TP156	35	M	NR	Portugal	NR	Várias	S	S	NR	N
TP158	34	F	NR	Portugal	NR	Portuguesa	S	N	A	N
TP159A	34	M	NR	Portugal	NR	Várias	S	S	NR	N
TP184A	35	M	NR	Angola	NR	Várias	S	S	NR	S

TP193	38	M	NR	Portugal	NR	Várias	S	S	NR	S
TP208	35	F	NR	Portugal	NR	Portuguesa	S	N	NR	N
TP211	31	M	NR	Portugal	NR	Portuguesa	S	S	NR	S
TP212	42	F	NR	Portugal	NR	Portuguesa	S	N	NR	N
TP219	36	M	NR	Moçambique	NR	Várias	S	N	NR	N
TX12	38	M	1994	Portugal	22000	NR	NR	S	A	NR
TX13	32	M	2005	Moçambique	27000	NR	NR	S	A	N
TX14	34	F	2005	Portugal	56000	NR	NR	S	A	N
TX15	47	M	2005	Portugal	90000	NR	NR	S	A	N
TX16	47	M	1989	NR	35000	NR	NR	S	A	NR
TX18	36	M	2006	NR	1900000	NR	NR	S	A	N
TX20	38	M	1993	NR	13000	NR	NR	S	A	NR
TX21	45	M	1996	NR	11000	NR	NR	S	A	NR
TX23	32	F	2000	NR	4200	NR	NR	S	A	NR
TX30	34	M	2006	NR	1200	NR	NR	S	A	N
TX32	41	F	1992	NR	3900	NR	NR	S	A	NR
TX37	24	M	2006	Portugal	160000	NR	NR	S	A	N
TX40	30	M	1995	Marrocos	190000	NR	NR	S	A	NR
TX41	30	F	2005	Portugal	13000	NR	NR	S	A	N
TX45	37	M	1998	Portugal	43000	NR	NR	S	A	NR
TX46	26	F	2004	Cabo Verde	16000	NR	NR	S	A	N
TX47	41	M	2005	Portugal	1700	NR	NR	S	A	N
TX48	34	M	1994	EUA	610	NR	NR	S	A	NR
TX50	36	M	1999	Portugal	140000	NR	NR	S	A	NR
TX59	29	M	1995	Portugal	11000	NR	NR	S	A	N
TX60	41	M	1996	Portugal	42000	NR	NR	S	A	NR
AS646	37	F	1998	Angola	NR	NR	NR	NR	NR	N

[ID: Identificação da amostra; M: Masculino; F: Feminino; NR: Não reportado; S: Sim; N: Não; A: Fase assintomática; CV: Carga Viral; TAR: Terapia anti-retroviral]

Anexo 3: Composição de meios e soluções

1. Meio LB:

10g Bacto-triptona, 5g extracto de levedura, 10g NaCl, 15g agar*.

Adicionar água para 1000ml e autoclavar.

*Não adicionar para LB líquido

2. Meio de criopreservação:

10% (v/v) DMSO, 10% (v/v) glicerol, em meio LB líquido.

3. Solução de agarose (1,5%) (p/v) (50ml):

0,75g agarose, 50ml TAE 0,5X.

4. Solução de agarose (1%) (p/v) (100ml):

1g agarose, 100ml TAE 0,5X.

5. Tampão Tris-acetato-EDTA (TAE) 50X:

242g Tris-base, 57,1ml ácido acético glacial, 100ml 0,5M EDTA, pH 8,0.

Adicionar água para 1000ml.

6. Tampão de aplicação:

0,25% (p/v) azul de bromofenol, 40% (p/v) sacarose.

7. Tampão da T4 DNA ligase 2X (Promega, EUA):

60mM Tris-HCl, pH 7,8, 20mM MgCl₂, 20mM DTT, 2mM ATP.

8. Meio de transformação e armazenamento (TSS):

LB 1X, 10% (v/v) PEG6000, 5% (v/v) DMSO, 50mM MgSO₄ (pH 6,5).

9. Placas LB agar/X-Gal/IPTG/ampicilina:

LB sólido [Bacto-agar a 1,5% (p/v) em LB líquido], 100µg/ml ampicilina, 40µg/ml X-Gal, 0,17mM IPTG.

10. Tampão Tris-EDTA-glicose (TEG):

50mM glicose, 10mM EDTA, 25mM Tris-HCl, pH 8,0.

11. Solução de Lise (SDS alcalino):

0,2 N NaOH, 1% (p/v) SDS.

12. Tampão da EcoRI 10X (Fermentas, Canadá):

50mM Tris-HCl, pH 7,5, 10mM MgCl₂, 100mM NaCl, 0,02% (v/v) Triton X-100, 0,1mg/ml BSA.